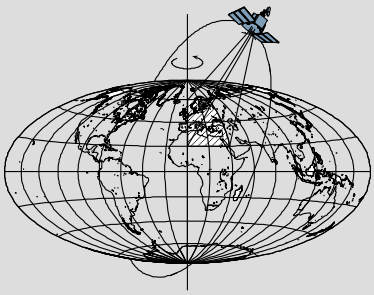
GRACE Time-Variable Gravity Field Recovery Using an Improved Energy Balance Formalism

by

Kun Shang



Report No. 511

Geodetic Science

The Ohio State University
Columbus, Ohio 43210

August 2015

GRACE Time-Variable Gravity Field Recovery Using an Improved Energy Balance Formalism

By

Kun Shang

Report No. 511

Geodetic Science

The Ohio State University
Columbus, Ohio 43210

August 2015

Preface

This Report was prepared for and submitted to the Graduate School of the Ohio State University as a dissertation in partial fulfillment of the requirements for the PhD degree.

This research is conducted under the supervision of Professor C.K. Shum, Division of Geodetic Science, School of Earth Sciences, The Ohio State University. This research is partially supported by grants from NSF via the Belmont Forum/IGFA Grant (ICER-1342644), and NASA's geodesy and cryosphere grants (NNX12AJ95G, NNX12AK28G, NNX11AR47G). GRACE data products are from NASA's PODAAC via Jet Propulsion Laboratory/California Institute of Technology (JPL), University of Texas Center for Space Research (CSR), and GeoForschungsZentrum Potsdam (GFZ). Some figures in this paper were generated using the Generic Mapping Tools (GMT) [*Wessel and Smith*, 1991]. The computational aspect of this work was supported in part by an allocation of computing resources from the Ohio Supercomputer Center (http://www.osc.edu).

Abstract

Earth's gravity is continuously varying with respect to time due primarily to mass transports within the Earth system and external gravitational forcing. A *new* formalism based on energy conservation principle for time-variable gravity field recovery using satellite gravimetry has been developed and yields more accurate estimation of in-situ geopotential difference observables using K-Band Ranging (KBR) measurements from the Gravity Recovery and Climate Experiment (GRACE) twin-satellite mission. The new approach can preserve more time-variable gravity information sensed by KBR range-rate measurements and reduce orbit error as compared to previous energy balance studies. Results based on analysis of more than 10 years of GRACE data indicate that the estimated geopotential differences agree well with the predicted values from official Level 2 solutions: with much higher correlation of 0.9, as compared to 0.5–0.8 reported by previous energy balance studies. This study demonstrates that the new approach is more flexible for both global and regional temporal gravity recovery, leading to the *first* independent GRACE monthly solution series based on energy conservation principle, which is comparable to the results from different approach. The developed formalism is applicable to the general case of low-low satellite-to-satellite radiometric or laser interferometric tracking measurements, such as GRACE Follow-on or other Next Generation Gravity Field missions, for efficient retrieval and studies of Earth's mass transport evolutions.

The regional gravity analysis over Greenland reveals that a substantially *higher* temporal resolution is achievable at 10 or 11-day interval from GRACE data, as compared to the official monthly solutions, but without the compromise of spatial resolution, nor the need to use regularization or post-processing. Studies of the terrestrial and ground water storage change over North China Plain show high correlation in sub-monthly scale, among the 11-day time-variable gravity solutions from this study, in-situ data, and hydrologic and atmospheric models. The 11-day solutions with 1-day step successfully capture the surface mass change caused by the rapid snow and ice accumulation and melting during the extreme weather event of 2008 Southeast China snow and ice storm. These results demonstrated that sub-monthly solutions from GRACE can provide an additional constraint to understand the rapid mass transport and the dynamic processes for both extreme weather events and short-time surface and ground water monitoring, which may potentially improve our understanding of various mass transports within the Earth system, and applicable to societal services such as disaster response or mitigation, and water resources management.

Table of Contents

Chapter 1 Introduction

1.1 GRACE Mission Overview

Launched in March 2002, the Gravity Recovery and Climate Experiment (GRACE) mission [*Tapley et al.*, 2004a] has been mapping Earth's time-variable gravity field from space at approximate 400~500 km altitude for more than a decade. The resulting remarkable scientific advances [*Cazenave and Chen*, 2010] have revolutionized our understanding both for solid earth study, such as Glacial Isostatic Adjustment [e.g., *Tamisiea et al.*, 2007; *van der Wal et al.*, 2015] and earthquakes [e.g., *Han et al.*, 2006; *Dai et al.*, 2014], and for mass transport studies on the Earth surface, such as ice sheet mass balance [e.g., *Velicogna et al.*, 2006; *Chen et al.*, 2006], oceanography [e.g., *Johnson et al.*, 2013], and hydrology [e.g., *Rodell et al.*, 2009; *Han et al.*, 2009]. From the data collected by the K/Ka-Band Ranging (KBR) low-low satellite-to-satellite tracking (SST) and the high-low GPS-GRACE twin-satellite tracking, monthly mean gravity field models in the form of Stokes coefficients (known as GRACE Level-2 products) have been routinely computed and made available publicly by the University of Texas Center for Space Research (CSR), GeoForschungsZentrum (GFZ) German Research Centre for Geosciences, Jet Propulsion Laboratory (JPL) and others. The estimation approach used by the three agencies and others [e.g., *Luthcke et al.*, 2006; *Bruinsma et al.*, 2010] to generate these solutions is the so-called dynamic method based on the dynamic orbit determination and parameter recovery principle [*Tapley et al.*, 2004b], i.e., a linearized least squares adjustment using observations for gravity coefficients, orbit and other parameters. Besides the conventional dynamic method, various similar or alternative approaches have also been proposed and implemented, such as mascon approach [*Rowland et al.*, 2005, 2010], short-arc approach [*Mayer-Gürr et al.*, 2007; *Kurtenbach et al.*, 2009], celestial mechanics approach [*Meyer et al.*, 2012], acceleration approach [*Chen et al.*, 2008; *Liu et al.*, 2010], and finally energy balance approach [*Jekeli*, 1999; *Visser et al.,* 2003; *Han et al.*, 2006; *Ramillien et al.*, 2011; *Tangdamrongsub et al.*, 2012]. The last approach is the focus of this study.

1.2 Energy Balance Approach

Energy balance approach, also known as energy integral approach, can be traced back to the 1960s (e.g., *Bjerhammer* [1967]) in the early era of satellite geodesy. The basic idea of this approach is to explore the possibility of applying the principle of energy conservation, i.e., the constant sum of kinematic energy and potential energy, to the satellite tracking data for direct measuring of Earth's gravity field. The concept was investigated again by *Jekeli* [1999] at the beginning of the Decade of Geopotential Missions, and developed the first practical formulation to explicitly express the

relationship between geopotential and satellite data in inertial frame (later called energy equation), with conceived application for the forthcoming satellite gravimetry missions, Challenging Minisatellite Payload (CHAMP) and GRACE. Shortly after, *Visser et al.* [2003] similarly derived the energy equation but in Earth-fixed frame. Since then, a renewed interest of using energy balance approach to estimate Earth's static and time-variable gravity field was aroused during the last decade, especially for applications using the data from satellite gravimetry missions, such as CHAMP [e.g., *Han et al.*, 2002; *Gerlach et al.*, 2003; *Badura et al.*, 2006], GRACE [e.g., *Han et al.*, 2006; *Ramillien et al.*, 2011; *Tangdamrongsub et al.*, 2012] and Gravity Field and Steady-State Ocean Circulation Explorer (GOCE) [e.g., *Pail et al.*, 2011].

One of the major advantages of energy balance formalism is that it is the sole approach that can be utilized to directly estimate the in-situ geopotential observables (for a single satellite) or geopotential difference observables (for a pair of satellites), which is realized through the so-called energy equation that represent a clear connection between the geometric measurements and geopotential. As a quantity with explicit geophysical interpretation, the in-situ geopotential observables are more natural to be used for gravity field inversions as compared with other approaches. The estimation procedure is also more efficient because of the linear relationship between the observables and gravity coefficients, that is, there is no linearization nor the assumption that the *a priori* state parameters have to be known sufficiently close to the true parameters. More importantly, the in-situ geopotential difference observables would greatly benefit the time-variable gravity recovery missions, such as GRACE, since the epoch-wise observables can support flexible spatial and temporal resolutions, leading to local enhanced solutions which could possibly retrieve more regional gravity information [*Han et al.*, 2005; *Schmidt et al.*, 2006, 2008; *Tangdamrongsub et al.*, 2012].

1.3 Motivation for This Study

However, appropriate applications of the energy balance approach to GRACE-type mission data for highly accurate geopotential estimation is still a demanding task. One of the most challenging problems is how to efficiently extract the gravity signal sensed by the essential measurements from SST, i.e., KBR range-rate measurements, of which the energy equation does not explicitly express. Previous researchers attempt to adjust range-rate and orbit data simultaneously via a nonlinear least squares estimation with either fixed constraints [*Han et al.*, 2006] or stochastic constraints [*Tangdamrongsub et al.*, 2012]. The use of constrained least squares adjustment, though straightforward, is still a compromise between the very high-precision range-rate data and the relative low-precision orbit data, which may tend to distort the estimation of in-situ geopotential observables caused by errors including orbit error. The orbit error, inherited from the chosen reference orbit, would contaminate the resulting gravity estimation especially at the low-frequency band [*Ditmar et al.*, 2012]. Besides, since orbit data are usually correlated with the *a priori* gravity model, any difference between the *a priori* models used in the orbit and the models used for energy method would cause systematic errors during the reduction of range-rate measurements. In addition, recent study [*Guo et al.*,

2015] has demonstrated that the previous formulation of the energy equation may contain a non-negligible approximation, which could overwhelm the time-variable gravity signal also at the low-frequency band. These issues limit the application of the earlier developed in-situ geopotential differences only to regional gravity analysis, i.e., at the high-frequency band, and arguably regions with large temporal gravity field signals. As a result, large-scale gravity field inversion, including global gravity solution, has not been fully exploited or were inaccurately applied based on previous energy approaches.

The primary purpose of this study is to overcome these limitations by employing an improved energy balance approach to obtain a more accurate estimation of in-situ geopotential difference observables, with an aim to preserve both the low- and high-frequency gravity signal and consequently yield a full scale, i.e., both regional and global, gravity inversion. To achieve this goal, I develop a novel formulation, called the alignment equation, to incorporate range-rate observations into energy equation, together with a method to reconstruct the related reference orbit. In addition, a more rigorous formulation of energy equation [*Guo et al.*, 2015] is applied to model the in-situ geopotential difference observables, which is proved to be requisite for the reduction of the GRACE measurement for gravity field inversion.

The objective of this study is also to use the resulting geopotential difference estimates to solve a comparable series of global monthly solution with official GRACE monthly products. Based on that, I aim to explore the possibility to improve the temporal resolution and recover any possible sub-monthly gravity change, and to address a few of the contemporary problems in Earth sciences. Three different cases are carried out to demonstrate the enhanced temporal resolution and the reliability of my gravity solutions.

1.4 Dissertation Outlines

The outline of this paper is as follows:

Chapter 2 presents the detailed methodology of the improved energy balance approach. It starts with a brief description of the general idea of energy balance approach formalism, following by the introducing of alignment equation and discussion about the essential advantage of this equation. After that the orbit reconstruction algorithm is given as well as the description of the *a priori* gravity models adopted in this study. The last section of this chapter describes the precise derivation of energy equation and the corresponding numerical results and their accuracy assessment.

Chapter 3 contains the results using the improved energy balance approach. The estimation of geopotential differences is given first as well as the calibration of empirical parameters for both accelerometer and range-rate. Next a new global monthly solution series based on energy conservation principle is presented from 2003 to 2013 using the geopotential difference estimates, and then the solution series are evaluated and compared with official GRACE monthly products for both the secular and seasonal gravity variation. Finally, I demonstrate that the solutions with enhanced temporal

resolutions are achievable and apply my sub-monthly solutions to study three contemporary Earth science problems. The first research topic focuses on the estimates of the Greenland ice sheet mass balance and ice mass evolutions. The second topic is to quantify terrestrial and ground water storage changes over the North China Plain aquifer using the improved GRACE solutions. The third topic is for the first detection of the 2008 Southeast China snow and ice storm using higher temporal sampled GRACE solutions.

Chapter 4 concludes the study and proposes future work.

Chapter 2 Improved Energy Balance Formalism

The orbit data of a near-Earth satellite, the positions and velocities within a well-defined reference frame, are dominated by the gravitational perturbations and other forces, and thus can be regarded as observations and used for gravity recovery, which is the basic concept of the energy balance approach. The energy equation, a mathematical expression of this concept, can be formulated in either Earth-fixed frame [*Visser et al*., 2003; *Zeng et al.,* 2015], or Earth-centered inertial frame [*Jekeli*, 1999]. The formulation in the Earth-centered inertial frame for a single satellite can be expressed as:

$$V = \frac{1}{2}|\dot{\mathbf{r}}|^2 + \int_{t_0}^{t} \frac{\partial V}{\partial t} dt - \int_{t_0}^{t} \mathbf{f} \cdot \dot{\mathbf{r}}\, dt - E^0, \tag{2.1}$$

where V is the total gravitational potential (for unit mass), $\mathbf{r}$ (implicit in V) and $\dot{\mathbf{r}}$ are the orbit position and velocity vectors in inertial frame, $\mathbf{f}$ is the non-conservative force, $\int_{t_0}^{t} (\partial V/\partial t)\, dt$ is the so-called potential rotation term due to time variations of the gravity field with respect to the inertial frame, which is mainly caused by the variable Earth rotation, and E^0 is a integral constant.

The total gravitational potential V can be decomposed into two parts $V=V^E+V^R$, where V^E is the geopotential, including both the Earth's mean, including secular, seasonal and other variable components which can be treated as a constant field during a short time, and V^R is the residual gravitational potential, mostly from the high-frequency (e.g., semi-diurnal and diurnal) varying geopotential, such as tides and high-frequency barotropic variation from atmosphere and ocean, which are assumed that they are known and will be removed from the observations. If we assume the residual gravitational potential V^R can be reduced or corrected using *a priori* models, and that the nonconservative force $\mathbf{f}$ can be measured by an onboard 3-axis accelerometer, we arrive at the complete formulation of energy equation for estimating geopotential V^E, from a single satellite, which can be expressed as

$$V^E = \frac{1}{2}|\dot{\mathbf{r}}|^2 + \int_{t_0}^{t} \frac{\partial V}{\partial t} dt - \int_{t_0}^{t} \mathbf{f} \cdot \dot{\mathbf{r}}\, dt - V^R - E^0. \tag{2.2}$$

And for estimating geopotential difference from a pair of satellites, such as GRACE, the formulation is simply the subtraction between the equations of two single satellites:

$$V_{12}^E = V_2^E - V_1^E = \frac{1}{2}|\dot{\mathbf{r}}_{12}|^2 + \dot{\mathbf{r}}_1 \cdot \dot{\mathbf{r}}_{12} + \int_{t_0}^{t} \frac{\partial V_{12}}{\partial t} dt - \int_{t_0}^{t} (\mathbf{f}_2 \cdot \dot{\mathbf{r}}_2 - \mathbf{f}_1 \cdot \dot{\mathbf{r}}_1) dt - V_{12}^R - E_{12}^0, \tag{2.3}$$

where the subscripts 1 and 2 represent the two satellites, and '12' their difference. The orbit data ($\mathbf{r}$ and $\dot{\mathbf{r}}$) in both formulations are normally regarded as observables, which can be obtained from a pre-computed reference orbit using high-low GPS tracking data.

Equations (2.2) and (2.3) are the widely used energy equations for the case of CHAMP and GRACE, respectively. The resulting geopotential or geopotential difference estimates at each orbital sampling point can be directly used for both global and regional gravity field recovery.

2.1 A Novel Method to Utilize Range-rate Measurements

2.1.1 Motivation of Seeking a New Method

Application of the energy balance approach on GRACE-type mission could be much more challenging than CHAMP-type mission partly because energy equation (2.3) is unable to explicitly contain the tracking measurements from the low-low SST system, i.e., range-rate measurements from KBR system in the case of GRACE. Previous studies usually use either of two methods to combine range-rate with the energy equation. One is an approximate method proposed by *Wolff* [1969], which relates the range-rate measurements to geopotential differences through a simplified equation as $V_{12} \approx \dot{\rho}|\dot{r}_1|$ or $V_{12} \approx \dot{\rho}|\dot{r}_1|$. But as an approximate method, it has not yet been successfully applied, since it is not suitable for high-precision geopotential estimation [*Morrison*, 1970], except for simulation studies [*Jekeli*, 1999]. The other method, as I mentioned before, is to treat range-rate measurements as redundancy observations, and use them with the orbit data, by a nonlinear least squares adjustment, where the energy equation is treated as either a fixed [*Han et al*., 2006], or as a stochastic [*Tangdamrongsub et al*., 2012] constraint. The estimates would be the six inter-satellite orbit states and other empirical parameters. However, it is known that the uncertainty of GRACE orbits is around 1~2 cm in positions and 10~20 μm/s in velocities (only for dynamic orbit; for kinematic orbit the uncertainty in velocities is even worse) [*Kang et al.*, 2006], whereas the range-rate measurements have a much lower uncertainty of about 0.2 μm/s [*Loomis et al.*, 2012]. The use of constrained least squares adjustment may be able to extract some information from range-rate measurement, but it is still a compromise between high-precision data and low-precision data, as the (unknown) systematic error, e.g., from orbit errors, would inevitably affect the solved parameters, and subsequently bias the estimation of geopotential difference observables.

In this study I aim to develop a new, alternative method to adjust inter-satellite orbit state parameters using range-rate measurements. The rationale is based on a simple fact, that a *single* range-rate measurement cannot be sensitive to *all* the six independent intersatellite orbit components (three relative position components and three relative velocity components), especially for the case of GRACE where the precision of range-rate is about two orders of magnitude higher than that of inter-satellite orbit data. Therefore, I

need to first find the most sensitive parameters(s) to the range-rate measurement, and then seek an approach to adjust or align the corresponding parameters(s) using range-rate measurements.

According to the previous study by *Rowlands et al*. [2002], the six independent intersatellite orbit components, i.e., the baseline vector can be transformed into a local Cartesian coordinate system. The *xy* plane of the local system can be defined to be perpendicular to the position vector of either satellite or the midpoint of the two satellites. The *x* axis is pointing to the local east and *y* axis is pointing to the flight direction. The *z* axis is defined according to the right-hand rule. Further, the Cartesian vector (relative position and velocity) in the local coordinate system can be converted into spherical coordinates, i.e., magnitude, pitch, which is the angle the vector makes with *xy* plane, and yaw, which is the angle that the projection onto the *xy* plane makes with the X axis. Under the new coordinate system, simulation study by *Rowlands et al*. [2002] has shown that, among all the inter-satellite parameters, the relative velocity pitch is the most sensitive parameter to range-rate measurements, and also one of the most important parameters for gravity recovery [*Luthcke et al*., 2006]. The other two important parameters are the relative velocity magnitude and the relative position pitch, but they are relatively much less sensitive to range-rate measurements as compared to the relative velocity pitch. Based on that, I develop a new equation to use range-rate measurements to adjust only the relative velocity pitch, i.e., the most sensitive parameter, and adopt or fix the other less sensitive or insensitive parameters as provided by the reference orbits.

2.1.2 Alignment Equation

The derivation of the new equation is straightforward. I start with the relative position vector, which can be written as $\mathbf{r}_{12} = \rho \mathbf{n}_1$, where $\mathbf{n}_1 = \mathbf{r}_{12} / |\mathbf{r}_{12}|$ is the unit vector along line-of-sight (LOS) direction, and ρ represent the relative range. Then I take derivative on the both sides, and get $\dot{\mathbf{r}}_{12}^{\rho} = \dot{\rho}\mathbf{n}_1 + \sqrt{|\dot{\mathbf{r}}_{12}|^2 - \dot{\rho}^2}\,\mathbf{n}_2$, where $\mathbf{n}_2$ is the unit vector of the derivative of $\mathbf{n}_1$. Since $\mathbf{n}_2$ have to be orthogonal to $\mathbf{n}_1$, $\mathbf{n}_2$ can be written as $\mathbf{n}_2 = (\mathbf{r}_{12} \times \dot{\mathbf{r}}_{12} \times \mathbf{r}_{12}) / |\mathbf{r}_{12} \times \dot{\mathbf{r}}_{12} \times \mathbf{r}_{12}|$. The projection along the direction $\mathbf{n}_2$ must be chosen as $\sqrt{|\dot{\mathbf{r}}_{12}|^2 - \dot{\rho}^2}$ in order to maintain the whole magnitude of the inter-satellite velocity as $|\dot{\mathbf{r}}_{12}|$.

Finally, I arrive at the new equation as follows:

$$\dot{\mathbf{r}}_{12}^{\dot{\rho}} = \dot{\rho}\frac{\mathbf{r}_{12}}{|\mathbf{r}_{12}|} + \sqrt{|\dot{\mathbf{r}}_{12}|^2 - \dot{\rho}^2}\,\frac{\mathbf{r}_{12} \times \dot{\mathbf{r}}_{12} \times \mathbf{r}_{12}}{|\mathbf{r}_{12} \times \dot{\mathbf{r}}_{12} \times \mathbf{r}_{12}|}, \tag{2.4}$$

which is referred to as the *alignment equation* throughout this paper. Here $\dot{\rho}$ represents the range-rate measurement, $\mathbf{r}_{12}$ and $\dot{\mathbf{r}}_{12}$ are the relative position and velocity vectors

from the reference orbit, $\mathbf{r}_{12} \times \dot{\mathbf{r}}_{12} \times \mathbf{r}_{12}$ is the vector triple cross product between them, and $\dot{\mathbf{r}}_{12}^{\dot{\rho}}$ represents the new relative velocity vector. As we can see, the new relative velocity vector $\dot{\mathbf{r}}_{12}^{\dot{\rho}}$ would be equal to the original vector $\dot{\mathbf{r}}_{12}$ if there is no additional range-rate observation, i.e., $\dot{\rho}$ is given as $\dot{\rho} = \dot{\mathbf{r}}_{12} \cdot \mathbf{r}_{12} / |\mathbf{r}_{12}|$. In that case, the alignment equation would degrade to an identical equation, which represents an exact geometric relationship between relative velocity direction vector and range-rate measurement. That means the alignment equation itself does not contain any approximation.

Once we have an independent and more accurate measuring of relative range-rate, such as the case of GRACE, then the new relative velocity vector $\dot{\mathbf{r}}_{12}^{\dot{\rho}}$ would become more accurate compared to the original vector $\dot{\mathbf{r}}_{12}$ because the pitch angle of relative velocity vector has been constrained by, or we can say, aligned to the range-rate. The term 'alignment' actually means the relative velocity pitch, the most sensitive inter-satellite parameter to range-rate and most important parameter for gravity recovery, has been aligned to the range-rate measurement through the equation. The reason can be further explained as follows.

The alignment equation essentially decomposes relative velocity vector into two components. One is along the line-of-sight (LOS) direction, where unit vector is $\mathbf{n}_1 = \mathbf{r}_{12} / |\mathbf{r}_{12}|$, and the correspondent projection is $\dot{\rho}$. The other is orthogonal to the LOS direction and is in the plane containing relative position vector and velocity vector, where the unit vector is $\mathbf{n}_2 = (\mathbf{r}_{12} \times \dot{\mathbf{r}}_{12} \times \mathbf{r}_{12}) / |\mathbf{r}_{12} \times \dot{\mathbf{r}}_{12} \times \mathbf{r}_{12}|$, with the correspondent projection of $\sqrt{|\dot{\mathbf{r}}_{12}|^2 - \dot{\rho}^2}$. In Figure 2.1, I illustrate such decomposition by showing the simple geometric configuration of GRACE constellation. As shown in Figure 2.1(b), the inter-satellite velocity is decomposed into two orthogonal directions along $\mathbf{n}_1$ and $\mathbf{n}_2$. The projection along the direction $\mathbf{n}_1$ should be equal to range-rate because of the relationship $\dot{\rho} = \dot{\mathbf{r}}_{12} \cdot \mathbf{r}_{12} / |\mathbf{r}_{12}|$ again. Then the other projection along the direction $\mathbf{n}_2$ should be equal to $\sqrt{|\dot{\mathbf{r}}_{12}|^2 - \dot{\rho}^2}$ in order to maintain the same magnitude of the inter-satellite velocity.

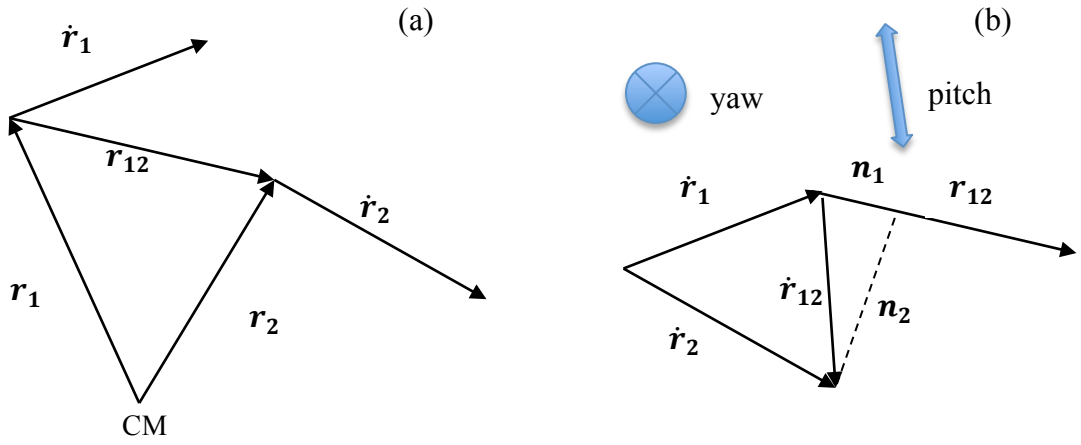


Figure 2.1 Geometric configuration of GRACE constellation and its relationship with alignment equation. (a) The absolute position and velocity vector for GRACE satellites with respect to the Earth's Center of Mass (CM). (b) The inter-satellite components of position and velocity vector and the decomposition of velocity vector, illustrating the derivation of the alignment equation.

Again, my goal here is to use high accurate measurement, i.e., the range-rate measured by the KBR, to adjust the most sensitive inter-satellite parameter, i.e., the relative velocity pitch. In another words, range-rate should be used to replace the relative velocity pitch component and form a new 'pitch-free' relative velocity vector. That is exactly what the alignment equation (2.4) represents. Under the decomposition as equation (2.4), the computation of $\dot{\mathbf{r}}_{12}^{\dot{\rho}}$ only requires four inter-satellite quantities, which are range-rate $\dot{\rho}$, relative velocity magnitude $|\dot{\mathbf{r}}_{12}|$, LOS direction unit vector $\mathbf{n}_1$, and direction unit vector $\mathbf{n}_2$ that is always perpendicular to $\mathbf{n}_1$ and in the plane of intersatellite position and velocity. Among the four quantities, two of them, $\dot{\rho}$ and $\mathbf{n}_1$, are totally independent of the velocity component (as well as the position magnitude), and $|\dot{\mathbf{r}}_{12}|$ is only dependent on the velocity magnitude. The last one, unit vector $\mathbf{n}_2$, does rely on the velocity direction, but only the yaw angle. Therefore, $\mathbf{n}_2$ does not depend on the relative velocity pitch at all. This can be further clarified using Figure 2.1 (b). Using the local spherical coordinate system, the error of the relative velocity vector can be decomposed into magnitude error, pitch error and yaw error. The original vector $\dot{\mathbf{r}}_{12}$ obviously contains all three kinds of error. In contrast, the new vector $\dot{\mathbf{r}}_{12}^{\dot{\rho}}$ by applying the alignment equation can totally eliminate pitch and its error of $\dot{\mathbf{r}}_{12}$. It should be emphasized that in Figure 2.1 (b) the four vectors, $\dot{\mathbf{r}}_1$, $\dot{\mathbf{r}}_2$, $\mathbf{r}_{12}$ and $\dot{\mathbf{r}}_{12}$, are not strictly in the same plane. The yaw error of $\dot{\mathbf{r}}_{12}$ can be explained as the uncertainty between the plane of $\mathbf{r}_{12}$ and $\dot{\mathbf{r}}_{12}$, and the plane of $\dot{\mathbf{r}}_1$ and $\dot{\mathbf{r}}_2$. Therefore, by using the alignment equation, the information that are needed from the reference orbit are only relative velocity magnitude, relative position direction, and relative velocity yaw, but not relative velocity pitch. Therefore, the resulting relative velocity vector $\dot{\mathbf{r}}_{12}^{\dot{\rho}}$ is a 'pitch-free' quantity. The effect of relative velocity pitch from

less accurate reference orbit is thus totally eliminated. The only contribution of the resulting relative velocity pitch is from range-rate, and therefore the most sensitive component to intersatellite observation, has been fully constrained by range-rate measurement through the alignment equation.

Also, previous publications usually adopted an equation $\dot{\rho} = \dot{\mathbf{r}}_{12} \cdot \mathbf{r}_{12} / |\mathbf{r}_{12}|$, such as equation (20) from *Jekeli* [1999] and equation (2) from *Han et al.* [2006], for GRACE data processing. Here I call it range-rate equation. Comparing the range-rate equation and alignment equation, we can see range-rate equation only represents the first component of the alignment equation, i.e., the projection along LOS direction. The alignment equation provides the other component, which is the direction that is orthogonal to the LOS direction, as well as the corresponding projection. By giving the full components, the alignment equation can successfully update the relative velocity vector using range-rate without any approximation or adjustment. For comparison, previous applications using range-rate equation, since it only has part of the components, can only consider it as a constraint during the procedure of least squares adjustment.

After applying the alignment equation, the new relative velocity vector $\dot{\mathbf{r}}_{12}^{\dot{\rho}}$, would subsequently be used as the input to the energy equation (2.3). Again, rather than merely providing a method to explicitly incorporate range-rate data into energy equation, alignment equation concentrates on aligning the most sensitive inter-satellite parameter using range-rate data, instead of equally adjusting it with all the other parameters. In the next step, I expect to determine geopotential difference observables solely from range-rate measurement, and meanwhile minimize the direct effect from the reference orbit to the estimates, which will be discussed in the next subsection.

2.2 Reconstruction of the Reference Orbit

The reference orbit, i.e., Cartesian coordinates from both satellites, is another critical input to the energy equation (2.3) in addition to the range-rate measurements for GRACE. In fact, in the case of CHAMP-type mission, since there are no additional geometric measurements, orbit data would be the only geometric input to the energy equation (2.2), besides accelerometer measurements. As we know, satellite can be regarded as a free-falling object that is falling under the sole influence of gravity field if non-conservative force can be neglected or completely accounted for. If one could precisely measure the orbit, both positions and velocities, thus the gravity information can be directly inferred from orbit data itself.

Therefore, satellite gravimetry highly relies on the orbit data instead of direct measurement from accelerometer or gravimeter. In fact, gravimetry mission like GRAIL (Gravity Recovery and Interior Laboratory) [*Konopliv et al.*, 2013] does not even carry an accelerometer or gravimeter, but the (lunar) gravity field can still be recovered by satellite-to-satellite and precise orbit tracking with the adequately accurate modeling of non-conservative forces.

2.2.1 Which Orbit?

2.2.1.1 Coupled Problem Between Orbit and Gravity

However, one must be very cautious about choosing the orbit data as the input of energy equation, because orbit products are usually coupled with gravity product for satellite gravimetry. Before the GPS era, all the satellite orbits, both on Earth and on other planets or the Moon, are dynamically computed using the *a priori* gravity field and other models. Those dynamic orbit data are merely byproduct of the gravity field and cannot be regarded as an observable. Using such orbit in energy method for the purpose of solving gravity would be problematic or even impossible since the orbit data does not contain any new gravity information. In the age of GPS, it becomes possible to geometrically compute or measure the satellite orbit on Earth, and these orbits would be less dependent of *a priori* gravity models as compared to the dynamic orbit.

Generally there exist three different choices of reference orbits, namely the kinematic, reduced-dynamic or dynamic orbit. For CHAMP, kinematic orbit is normally preferred to dynamic or reduced-dynamic orbit because arguably the kinematic orbit is relatively free of *a priori* gravity models. The estimation of geopotential using kinematic orbit could avoid being biased toward the *a priori* gravity model, but the drawback is that a special processing, such as interpolation, has to be applied in order to obtain the velocity components since kinematic orbits usually can only provide position components, which normally would introduce large interpolation error to velocity components in kinematic orbit [*Gerlach et al.*, 2003].

The situation is different for the case of GRACE since there exists a redundant and more precise geometric observation, i.e., the KBR range-rate measurement. Range-rate should dominate the time-variable gravity information, so the geopotential difference observables do not rely much on the choice of the reference orbit. Therefore, it is possible to choose dynamic or reduced-dynamic orbit as the reference orbit for GRACE gravity field recovery, as long as the range-rate measurements are appropriately used to correct or adjust the orbit data. So in practice, various reference orbit data have indeed been implemented for GRACE real data analysis [*Han et al.*, 2006; *Tangdamrongsub et al.*, 2012].

It is worth mentioning that directly using those pre-computed dynamic or reduced dynamic orbits requires the user must fully comprehend all the *a priori* models as well as the parameterization for the orbit determination, and otherwise any difference between the *a priori* models used in the orbit and the referenced models used for energy method would cause systematic errors in the reduction of KBR measurements. On the other side, directly using kinematic orbit could avoid these systematic errors, but still would suffer from the random error of the kinematic position as well as the interpolation error of the velocity components.

2.2.1.2 Reconstruction of Purely Dynamic Orbit

In order to avoid any mis-modeled or un-modeled error from the pre-computed dynamic or reduce-dynamic orbit, and suppress the random and interpolation error from the kinematic orbit, I choose to adopt a purely dynamic orbit as the reference orbit in this study. The purpose also aims to partly exclude the direct contribution from the low-precision reference orbit (compared to range-rate) to the geopotential difference estimates. As I said before, when a purely dynamic orbit computed from *a priori* gravity model is used as the only input for energy equation, the output from both equation (2.2) and (2.3) would be inevitably reduced to the same *a priori* gravity model. That means no new geopotential information would be obtained from energy equation if only the purely dynamic orbits are used without range-rate measurements. Accordingly, once the purely dynamic orbit has been aligned through the alignment equation by including the range-rate measurements, the updated orbit would contain the new time-variable gravity information propagated solely from the rang-rate measurement, which would be revealed by the energy equation afterwards. By this means, all the new gravity information comes from the discrepancy between the range-rate measurement and the dynamic orbit computed from *a priori* gravity model. If the method can be iteratively applied, i.e., using the solved gravity solution as a new *a priori* gravity model, the discrepancy would be gradually reduced after each iteration, and the iterated gravity solution would be closer to the true gravity field.

The drawback for this method (and probably for all other implements based on energy balance principle so far) is that the signal-coherent adjustment is inevitable. The range-rate measurement (and also the accelerometer data) does not only contain signals and white noise, but also systematic error mainly driven by the orbital revolution. These systematic errors have to be adjusted with the orbit data, meaning the solved systematic parameters are always biased to the *a priori* gravity model, which is a kind of signal-coherent adjustment. Again iteration might be able to overcome this problem however it will be a topic of the future work.

Nevertheless, comparing to previous studies based on pre-computed orbit, I claim that the use of a purely dynamic orbit, together with the new method as alignment equation to incorporate the range-rate measurements, would have positive impact on the accuracy of the resulting GRACE disturbance potential difference observables.

Instead of computing a dynamic orbit directly from GPS observation, I use an alternative, but simple method to reconstruct the purely dynamic orbit from existing orbit data products. The similar technique has been used for previous studies on GRACE [*Liu et al.*, 2010] and GOCE [*Yi*, 2012]. The idea is to treat the available orbit coordinates (positions and/or velocities) as pseudo observations, and estimate a purely dynamic orbit by fitting the orbit coordinates with respect to a complete reference model, via least squares adjustment. Meanwhile, the accelerometer data are also simultaneously calibrated with respect to the purely dynamic orbit. The detailed algorithm for the orbit reconstruction is given in next section.

2.2.2 Orbit Reconstruction Algorithm

The orbit reconstruction algorithm is based on the principle of statistical orbit determination [*Tapley et al.*, 2004b] and can be briefly summarized as follows.

2.2.2.1 Observation Equation and Linearization

First, the observation equation, i.e., observation-state relationship, can be expressed as

$$O = F(\Xi) + e\,, \tag{2.5}$$

where O denotes the vector of observations, Ξ denotes the unknown state vector to be estimated, and e denotes the vector of unknown observation errors. For this particular problem, O represents the vector of pre-computed orbit coordinates at each epoch and Ξ represents the reconstructed orbit coordinates at each epoch and piecewise accelerometer parameters.

Observation O and unknown states Ξ have a significant nonlinear relationship F. Therefore, either a nonlinear adjustment should be applied, or the linearization has to be conducted first, followed by a linear adjustment and iteration. In the general orbit determination problem, the latter is a routine procedure. So I linearize the equation (2.5) by applying Taylor-series expansion about a reference trajectory and neglecting terms of order higher than the first as

$$O \approx F(\Xi_0) + \left.\frac{\partial F}{\partial \Xi}\right|_{\Xi=\Xi_0} (\Xi - \Xi_0) + e\,, \tag{2.6}$$

where Ξ_0 is the reference state vector of the unknowns. So the referenced observation vector $C = F(\Xi_0)$ can be computed from the reference trajectory and the nonlinear relationship F. By defining the residual observation as $y = O - C = O - F(\Xi_0)$, the partial derivative matrix as $A = \left.\frac{\partial F}{\partial \Xi}\right|_{\Xi=\Xi_0}$, and the residual unknown state vector as $\xi = \Xi - \Xi_0$, I can rewrite the linearized equation (2.6) as

$$y = A\xi + e\,, \tag{2.7}$$

which can be regarded as a linearized Gauss-Markov Model (GMM) if observation errors are considered to be random and have zero expectation as $e \sim (0, \sigma^2 P^{-1})$. So the LEast-Squares Solution (LESS) of this GMM can be given by the expressions

$$\hat{\xi} = \left(A^T PA\right)^{-1} A^T Py\ ,\ D\left(\hat{\xi}\right) = \sigma^2 \left(A^T PA\right)^{-1}. \tag{2.8}$$

Therefore the estimates for the unknown states and the associated dispersion and covariance matrix can be written as

$$\hat{\Xi} = \Xi_0 + \hat{\xi}\ ,\ D\left(\hat{\Xi}\right) = D\left(\hat{\xi}\right).$$

2.2.2.2 Partial Derivative Matrix

The partial derivative matrix A of nonlinear function F with respect to unknown states Ξ is the only term that still needs to be derived. Proper categorization of the unknown states will help to clarify the partial derivative formulation. The unknown states Ξ can be divided into two groups as dynamic parameters and kinematic parameters. Dynamic parameters are the orbit-related parameter $X(t)$, i.e., position and velocity vectors, which vary with respect to time due to the dynamic process and can be mapped into other states by using a state transition matrix. Kinematic parameters are the additional parameters B, such as accelerometer parameters here, which are usually regarded as constant during the adjustment. So the unknown states can be written as $\Xi = \left[\mathrm{B}, X(t)\right]$. After linearization, I have

$$\xi = \left[\beta, x(t)\right]. \tag{2.9}$$

Here I use uppercases to represent the states and parameters and the lowercase to represent the residual of states and parameters. The dynamic parameters $x(t)$ are not independent at each epoch because of the dynamic process, and can be propagated from the initial dynamic parameters $x(t_0)$ and kinematic parameters by a linear system called state transition equation as

$$x(t) = \left.\frac{\partial X(t)}{\partial X(t_0)}\right|_{X_0(t_0)} x(t_0) + \left.\frac{\partial X(t)}{\partial \mathrm{B}}\right|_{\mathrm{B}_0} \beta. \tag{2.10}$$

Therefore, the number of the unknown states can be reduced using the initial state as

$$\xi = \left[\beta, x(t_0)\right], \tag{2.11}$$

i.e., converting a boundary value problem (2.9) of a differential equation into an initial value problem (2.11), which is the basic concept of dynamic orbit determination.

Then I expand the previous GMM equation (2.7) as

$$
\begin{aligned}
y = A\xi + e &= \left.\frac{\partial F}{\partial \Xi}\right|_{\Xi=\Xi_0} \xi + e \\
&= \left.\frac{\partial F}{\partial X(t)}\right|_{X_0(t)} x(t) + \left.\frac{\partial F}{\partial \mathrm{B}}\right|_{\mathrm{B}_0} \beta + e
\end{aligned}
\quad . \tag{2.12}
$$

Substituting the state transition equation (2.10) leads to

$$
\begin{aligned}
& y = A\xi + e \\
&= \left.\frac{\partial F}{\partial X(t)}\right|_{X_0(t)} \left(\left.\frac{\partial X(t)}{\partial X(t_0)}\right|_{X_0(t_0)} x(t_0) + \left.\frac{\partial X(t)}{\partial \mathrm{B}}\right|_{\mathrm{B}_0} \beta \right) + \left.\frac{\partial F}{\partial \mathrm{B}}\right|_{\mathrm{B}_0} \beta + e \\
&= \left(\left.\frac{\partial F}{\partial X(t)}\right|_{X_0(t)} \left.\frac{\partial X(t)}{\partial X(t_0)}\right|_{X_0(t_0)} \right) x(t_0) + \left(\left.\frac{\partial F}{\partial X(t)}\right|_{X_0(t)} \left.\frac{\partial X(t)}{\partial \mathrm{B}}\right|_{\mathrm{B}_0} + \left.\frac{\partial F}{\partial \mathrm{B}}\right|_{\mathrm{B}_0} \right) \beta + e
\end{aligned}
\quad . \tag{2.13}
$$

Hence I arrive at the expression of the partial derivative matrix *A*.

Here, partial derivatives $\left.\frac{\partial F}{\partial X(t)}\right|_{X_0(t)}$ and $\left.\frac{\partial F}{\partial \mathrm{B}}\right|_{\mathrm{B}_0}$ can be computed from the observation equation. For this particular problem of orbit reconstruction, the observation equation can be written as

$$
O = X(t, X_0, \mathrm{B}) + e \, . \tag{2.14}
$$

Therefore, through easy derivation, I get

$$
\left.\frac{\partial F}{\partial X(t)}\right|_{X_0(t)} = \left.\frac{\partial X}{\partial X(t)}\right|_{X_0(t)} = I \quad \text{and} \quad \left.\frac{\partial F}{\partial \mathrm{B}}\right|_{\mathrm{B}_0} = \left.\frac{\partial X}{\partial \mathrm{B}}\right|_{\mathrm{B}_0} = 0 \, . \tag{2.15}
$$

The other two partial derivatives $\left.\frac{\partial X(t)}{\partial X(t_0)}\right|_{X_0(t_0)}$ and $\left.\frac{\partial X(t)}{\partial \mathrm{B}}\right|_{\mathrm{B}_0}$, called state transition matrices, can be derived from solving state transition equation (2.10).

2.2.2.3 State Transition Matrix

I denote the two state transition matrices as Φ_1 and Φ_2, i.e., $\Phi_1 = \left.\frac{\partial X(t)}{\partial X(t_0)}\right|_{X_0(t_0)}$ and $\Phi_2 = \left.\frac{\partial X(t)}{\partial \mathrm{B}}\right|_{\mathrm{B}_0}$. Then the state transition equation (2.10) can be rewritten as

$$x(t) = \Phi_1 x(t_0) + \Phi_2 \beta \,. \tag{2.16}$$

Now, I need to consider the dynamic equations, which can be expressed as

$$\begin{aligned} \dot{X}(t) &= f\left(X(t), \mathrm{B}\right) \\ \dot{\mathrm{B}} &= 0 \end{aligned}, \tag{2.17}$$

where the first equation represents the equations of motion for the satellite, and the second equation represents the constant parameters. Similarly, after linearization I have

$$\begin{aligned} \dot{x}(t) &= \left.\frac{\partial \dot{X}(t)}{\partial X(t)}\right|_{X_0(t)} x(t) + \left.\frac{\partial \dot{X}(t)}{\partial \mathrm{B}}\right|_{\mathrm{B}_0} \beta \\ \dot{\beta} &= 0 \end{aligned}. \tag{2.18}$$

Next, by substituting state transition equation (2.16) to the first equation of (2.18), I have

$$\begin{aligned} \dot{x}(t) &= \left.\frac{\partial \dot{X}(t)}{\partial X(t)}\right|_{X_0(t)} \left(\Phi_1 x(t_0) + \Phi_2 \beta\right) + \left.\frac{\partial \dot{X}(t)}{\partial \mathrm{B}}\right|_{\mathrm{B}_0} \beta \\ &= \left(\left.\frac{\partial \dot{X}(t)}{\partial X(t)}\right|_{X_0(t)} \Phi_1\right) x(t_0) + \left(\left.\frac{\partial \dot{X}(t)}{\partial X(t)}\right|_{X_0(t)} \Phi_2 + \left.\frac{\partial \dot{X}(t)}{\partial \mathrm{B}}\right|_{\mathrm{B}_0}\right) \beta \end{aligned}. \tag{2.19}$$

On the other hand, by directly taking derivatives on both sides of state transition equation (2.16), I also have

$$\dot{x}(t) = \dot{\Phi}_1 x(t_0) + \dot{\Phi}_2 \beta \,. \tag{2.20}$$

Comparing the two equations above, (2.19) and (2.20), I arrive at

$$\begin{aligned}\dot{\Phi}_1 &= \left.\frac{\partial \dot{X}(t)}{\partial X(t)}\right|_{X_0(t)} \Phi_1 \\ \dot{\Phi}_2 &= \left.\frac{\partial \dot{X}(t)}{\partial X(t)}\right|_{X_0(t)} \Phi_2 + \left.\frac{\partial \dot{X}(t)}{\partial \mathrm{B}}\right|_{\mathrm{B}_0}\end{aligned}. \tag{2.21}$$

Equation (2.21) are the differential equations that the state transition matrices should satisfy, which means the state transition matrices can be solved by integrating equation (2.21) with a proper initial condition.

The initial conditions of the transition matrices can be found from equation (2.16) by letting $t = t_0$, which leads to

$$x(t_0) = \Phi_1(t_0)x(t_0) + \Phi_2(t_0)\beta\,. \tag{2.22}$$

That is, the initial conditions of the transition matrices can be expressed as

$$\begin{aligned}\Phi_1(t_0) &= I \\ \Phi_2(t_0) &= 0\end{aligned}. \tag{2.23}$$

Therefore, the problem of orbit reconstruction can be fully solved by numerically integrating differential equations of state transition matrices (2.21) and dynamic equations (2.17), in order to fulfill the partial derivative matrix A in the GMM (2.13). The only thing left is the two partial derivative matrices during the linearization of the dynamic equation, i.e., $\left.\frac{\partial \dot{X}(t)}{\partial X(t)}\right|_{X_0(t)}$ and $\left.\frac{\partial \dot{X}(t)}{\partial \mathrm{B}}\right|_{\mathrm{B}_0}$. In the next section, I will describe the background dynamic models used in this study and derive the two partial derivative matrices using the corresponding background models.

2.2.3 Dynamic Model and Input Orbit Data

2.2.3.1 Dynamic Equation and its Partial Derivative

As I said before, during the process of the orbit reconstruction, the dynamic models would serve two purposes. One is for the equations of motion on the right-hand side of the first equation of (2.17), which has to be as accurate as possible. The other is for the partial derivative matrix on the right-hand side of the first equation of (2.18) and further for the state transition matrix, which can be approximate since it is already a linearization of the dynamic equation. The general form of dynamic equation can be expressed as

$$\begin{aligned}
\dot{r}_x &= v_x \\
\dot{r}_y &= v_y \\
\dot{r}_z &= v_z \\
\dot{v}_x &= a_x \\
\dot{v}_y &= a_y \\
\dot{v}_z &= a_z
\end{aligned} \tag{2.24}$$

where $\mathbf{r} = \left[r_x, r_y, r_z \right]^T$ is the position vector, $\mathbf{v} = \left[v_x, v_y, v_z \right]^T$ is the velocity vector, and $\mathbf{a} = \left[a_x, a_y, a_z \right]^T$ is the acceleration vector. Using the above notation, that is

$$X(t) = \left[\mathbf{r}^T, \mathbf{v}^T \right]^T = \left[r_x, r_y, r_z, v_x, v_y, v_z \right]^T \tag{2.25}$$

$$\dot{X}(t) = \left[\mathbf{v}^T, \mathbf{a}^T \right]^T = \left[v_x, v_y, v_z, a_x, a_y, a_z \right]^T \tag{2.26}$$

The partial derivative matrix of the satellite state time derivative with respect to satellite state is given as

$$\frac{\partial \dot{X}(t)}{\partial X(t)} = \begin{bmatrix} 0_{3\times3} & I_{3\times3} \\ \left(\dfrac{\partial \mathbf{a}}{\partial \mathbf{r}} \right)_{3\times3} & \left(\dfrac{\partial \mathbf{a}}{\partial \mathbf{v}} \right)_{3\times3} \end{bmatrix} \tag{2.27}$$

which is a 6 by 6 matrix.

The partial derivative matrix of the satellite state time derivative with respect to constant parameters is given as

$$\left. \frac{\partial \dot{X}(t)}{\partial \mathrm{B}} \right|_{\mathrm{B}_0} = \begin{bmatrix} 0_{3\times m} \\ \left(\dfrac{\partial \mathbf{a}}{\partial \mathrm{B}} \right)_{3\times m} \end{bmatrix} \tag{2.28}$$

which is a 6 by m matrix, if the number of the constant parameters is m.

By substituting equation (2.27) and (2.28) into (2.21), one can further solve the differential equation of the state transition matrix. Including the 6 differential equations of motion, the total number of the differential equations that need to solve is $6 + 6 \times 6 + 6 \times m = 42 + 6m$. All the differential equations can be simultaneously solved

using a numerical integrator. The detailed dynamic equation of motion is given in next section.

2.2.3.2 Dynamic Models Used in This Study

The equation of motion can be explicitly expressed as

$$\ddot{\mathbf{r}} = \mathbf{a} = \mathbf{g} + \mathbf{f} + \mathbf{a}_{res} \tag{2.29}$$

where **g** represents all conservative forces, such as geopotential, tide, non-tidal variation, N-body perturbation and some other conservative forces, **f** represents non-conservative force measured by the on-board accelerometer and orientation data, $\mathbf{a}_{res}$ represents other residual accelerations, such as general relativity perturbation.

The background models adopted in this study are identical to the models used by GFZ for solving the official GRACE Level-2 (L2) product Release 05 (RL05) [*Dahle et al.*, 2012], which include geopotential model from EIGEN-6C up to degree and order 200, anelasticity solid Earth tides model with frequency dependent corrections from IERS 2010 Standard [*Petit and Luzum*, 2010] (Section 6.2), EOT11a ocean tides model from *Mayer-Gürr et al.* [2012], atmosphere and oceanic non-tidal variability from GRACE AOD1B RL05 product, pole tide model from IERS 2010 (Section 6.4), N-body perturbation with indirect J_2 effect from the DE421 ephemerides, general relativistic perturbation from IERS 2010 (Section 10.3), atmosphere tides from *Biancale and Bode* [2006], and ocean pole tide model from IERS 2010 (section 6.5). The accelerometer data from GRACE Level 1B (L1B) ACC1B product, together with orientation data from L1B SCA1B product, are used to model the non-gravitational acceleration.

2.2.3.3 Input Pre-computed Orbit Products

There are many existing orbit data products available to potentially serve as the input of the reconstruction of the purely dynamic orbit. I have tested three different highly accurate scientific orbit products from independent institutes, which are the kinematic orbit product from National Central University, Taiwan [Courtesy, *Tzupang Tseng*], the kinematic orbit product from University of Bern [Courtesy, *Adrian Jäggi*], and the reduced-dynamic orbit from JPL, i.e., the GRACE L1B GNV1B product [Courtesy, *Dahning Yuan*]. I found the difference between the resulting reconstructed purely dynamic orbits using these orbit products negligible. In addition, the subsequent geopotential difference observables are also not sensitive to the input orbit product because of the reconstruction process. Therefore, I simply choose the L1B GNV1B product as the input orbit data to reconstruct the purely dynamic orbit since it is readily available with other L1B products.

The reconstructed dynamic orbit is first used as the input of the alignment equation (2.4), together with the range-rate data from GRACE L1B KBR1B product. The resulting new

relative velocity vector $\dot{\mathbf{r}}_{12}^{\dot{\rho}}$ and the accelerometer calibration parameters from the reconstruction process are then used as the input data to the energy equation (2.3). Again, this process will minimize the direct effect from the reference orbit to the geopotential differences simply because the estimates would be reduced to the *a priori* gravity model if range-rate data were absent. New gravity information could be revealed only if the range-rate data is used to correct orbit data through the alignment equation. One may argue that this process may also eliminate the possible contribution from GPS tracking data to gravity estimation, but considering the much higher (~50 times) noise level (both high frequency and low frequency) in the orbit data as compared to the KBR range-rate data, I believe it is a reasonable tradeoff. Besides, it is worth pointing out this strategy is also used by *Luthcke et al.* [2006]. They applied the traditional dynamic method to solve GRACE monthly solution, but also solely from range-rate measurements, and claim they achieved similar or even better monthly solutions than the official GRACE Level 2 data products.

2.3 Formulation and Assessment of Energy Equation

When energy equation (2.3) is applied to compute the geopotential difference, all the quantities and terms on the right-hand side can be computed from data or based on reference models, except one term, $\int_{t_0}^{t}(\partial V/\partial t)dt$, the so-called "potential rotation term". The potential rotation term represents the rate of change of the total potential V in the geocentric inertial frame, and the integral is a line integral following the orbits of the two satellites from t_0 to t. In energy equations of both equation (2.2) for single satellite and equation (2.3) for the twin satellites, the potential rotation term must be reformulated for practical purpose; otherwise, the unknown, geopotential V^E would appear in both sides of the equation. This subsection will discuss how to formulate the energy equation. Here only the single satellite case (equation (2.2)) needs to be considered, since the twin satellite case (equation (2.3)) is merely the subtraction between the equations of two single satellites.

The formulation of energy equation (2.2) previously developed by *Jekeli* [1999] is as follows:

$$V^E \approx \frac{1}{2}|\dot{\mathbf{r}}|^2 - \omega(x\dot{y} - y\dot{x}) - \int_{t_0}^{t}\mathbf{f}\cdot\dot{\mathbf{r}}\,dt - V^R - E^0\,, \tag{2.30}$$

where the potential rotation term was approximated as $\int_{t_0}^{t}(\partial V/\partial t)dt \approx -\omega(x\dot{y} - y\dot{x})$, x and y represent the first and second component for position vector, and ω is the nominal Earth's angular velocity, i.e., Earth's angular velocity vector along the third axis.

However, *Guo et al.* [2015] have already demonstrated via simulation that this approximation of the potential rotation term does not fulfill the precision of GRACE

observation, because the contribution of the time variable part of the gravitational potential to the potential rotation term was partly neglected. Although *Ramillien et al.* [2011] computed the potential rotation term in a different way, the same approximation has been made. Instead, a more accurate formulation of the energy equation should be used as follows:

$$V^E \approx \frac{1}{2}|\dot{\mathbf{r}}|^2 - \mathbf{w}\cdot(\mathbf{r}\times\dot{\mathbf{r}}) - \int_{t_0}^{t} \mathbf{a}\cdot(\dot{\mathbf{r}} - \mathbf{w}\times\mathbf{r})\,dt - E^0\,, \tag{2.31}$$

where $\mathbf{a} = \nabla V^R + \mathbf{f}$ is the acceleration of both residual geopotential acceleration and nonconservative acceleration, and $\mathbf{w}$ is Earth's angular velocity of Earth-fixed frame relative to the inertial frame, with coordinates in the inertial frame. The third term of the right-hand side can be numerically integrated. Similar formulation can be also found in previous studies especially for CHAMP [e.g. *Badura et al.*, 2006; *Jäggi et al.*, 2008]. In the next section, I give the detailed derivation of the formulation (2.31).

2.3.1 Derivation of the Formulation

2.3.1.1 Newton's Law of Motion in Inertial and Earth-fixed Frame

The position vector $\mathbf{r}$, representing the satellite position relative to Earth center, can be expanded in any Cartesian coordinate frame *s* (called *s*-frame) with a set of unit vectors denoted by $\mathbf{e}_j^s$, j=1, 2, 3. Using the unit vector, this position vector can be further rewritten as a *coordinate vector* $\mathbf{r}^s = r_1^s\mathbf{e}_1^s + r_2^s\mathbf{e}_2^s + r_3^s\mathbf{e}_3^s$. Or it can be represented as an ordered triplet of coordinates in s-frame as $\mathbf{r}^s = \begin{pmatrix} r_1^s & r_2^s & r_3^s \end{pmatrix}^T$. Two Earth-centered coordinate frames will be primarily considered, the inertial frame or *i*-frame defined as a frame where Newton's law of motion holds, and the rotational Earth-fixed frame or *e*-frame. Therefore, the position vector can be projected into both frames, denoted by coordinate vector $\mathbf{r}^i$ in *i*-frame, and coordinate vector $\mathbf{r}^e$ in Earth-fixed frame (*e*-frame), respectively. The relation between the two coordinate vectors can be described as

$$\mathbf{r}^i = C_e^i\mathbf{r}^e\,, \tag{2.32}$$

where C_e^i is a transformation matrix representing the orientation between the two frames. The time-derivative of C_e^i can be derived using the angular velocity $\mathbf{w}$ between the two frames. Here I define $\mathbf{w}_{ie}^i = (\omega_1, \omega_2, \omega_3)^T$ as the angular velocity vector of the *e*-frame with respect to the *i*-frame, with coordinates in the *i*-frame. The cross product of the angular velocity vector can be further written as a skew-symmetric matrix

$$\left[\mathbf{w}_{ie}^{i}\times\right]=\Omega_{ie}^{i}=\begin{pmatrix}0 & -\omega_3 & \omega_2\\ \omega_3 & 0 & -\omega_1\\ -\omega_2 & \omega_1 & 0\end{pmatrix}.$$

Using above notations, the time-derivative of the transformation matrix C_e^i can be derived as [*Jekeli*, 2001]

$$\dot{C}_e^i=-\Omega_{ei}^i C_e^i=\Omega_{ie}^i C_e^i=C_e^i\Omega_{ie}^e=\left[\mathbf{w}_{ie}^i\times\right]C_e^i=C_e^i\left[\mathbf{w}_{ie}^e\times\right].$$

Therefore, taking the time-derivative of equation (2.32) yields

$$\dot{\mathbf{r}}^i=C_e^i\dot{\mathbf{r}}^e+\dot{C}_e^i\mathbf{r}^e=C_e^i\left(\dot{\mathbf{r}}^e+\mathbf{w}_{ie}^e\times\mathbf{r}^e\right), \tag{2.33}$$

which represent the transformation of the velocity vector between two frames. Taking another time-derivative of equation (2.33) yields

$$\ddot{\mathbf{r}}^i=C_e^i\ddot{\mathbf{r}}^e+2C_e^i\left(\mathbf{w}_{ie}^e\times\dot{\mathbf{r}}^e\right)+C_e^i\left[\mathbf{w}_{ie}^e\times\left(\mathbf{w}_{ie}^e\times\mathbf{r}^e\right)\right], \tag{2.34}$$

which represents the transformation of the acceleration vector between two frames. Here the assumption $\dot{\mathbf{w}}_{ie}=\dot{\mathbf{w}}_{ie}^i=\dot{\mathbf{w}}_{ie}^e=0$ is required, which is the first assumption for energy equations.

Acceleration vector in *i*-frame must obey Newton's law of motion, which says

$$\ddot{\mathbf{r}}^i=\nabla^i V^E+\mathbf{a}^i, \tag{2.35}$$

where $\nabla^i=\begin{pmatrix}\frac{\partial}{\partial r_1^i} & \frac{\partial}{\partial r_2^i} & \frac{\partial}{\partial r_3^i}\end{pmatrix}^T$ represents the gradient operator in *i*-frame, V^E is Earth's static as well as secular and seasonal time-variable geopotential, and $\mathbf{a}^i=\nabla^i V^R+\mathbf{f}^i$ represent the sum of the acceleration from residual geopotential V^R and nonconservative force **f**.

In *e*-frame, Newton's law of motion can be derived using equation (2.34) and (2.35) as

$$\ddot{\mathbf{r}}^e=\nabla^e V^E+\mathbf{a}^e-2\left(\dot{\mathbf{w}}_{ie}^e\times\dot{\mathbf{r}}^e\right)-\left[\dot{\mathbf{w}}_{ie}^e\times\left(\dot{\mathbf{w}}_{ie}^e\times\mathbf{r}^e\right)\right], \tag{2.36}$$

where all quantities are represented in *e*-frame.

2.3.1.2 Time-Derivative of Static Geopotential in Both Frames

For the purpose of derivation of energy equation, the time-derivative of Earth's static gravitational potential V_E is also needed, which can be conducted in either *i*-frame or *e*-frame. In *i*-frame, the time-derivative of V^E can be expanded as

$$\dot{V}^E = \frac{dV^E}{dt} = \left.\frac{\partial V^E}{\partial t}\right|^i + \left(\nabla^i V^E\right)^T \cdot \dot{\mathbf{r}}^i, \tag{2.37}$$

The same is for *e*-frame where the time-derivative of V^E can be expanded as

$$\dot{V}^E = \frac{dV^E}{dt} = \left.\frac{\partial V^E}{\partial t}\right|^e + \left(\nabla^e V^E\right)^T \cdot \dot{\mathbf{r}}^e, \tag{2.38}$$

In equation (2.37) and (2.38) V^E appears on both side of each equation. The goal next is to rewrite both equations by eliminating V^E from the right-hand side. Here the second assumption $\left(\partial V^E/\partial t\right)\big|^e = 0$ is needed, which means V^E has to be static in *e*-frame, i.e., the partial derivative with respect to time in *e*-frame should be zero, for a certain time interval. So by combining equation (2.37) and (2.38) under the second assumption, I get

$$\left.\frac{\partial V^E}{\partial t}\right|^i = \left(\nabla^i V^E\right)^T \left(C_e^i \dot{\mathbf{r}}^e - \dot{\mathbf{r}}^i\right).$$

Substituting the above equation into (2.37) by considering (2.33) and Newton's law in *i*-frame (2.35), I can rewrite equation (2.37), the time-derivative of static geopotential in *i*-frame, as

$$\frac{dV^E}{dt} = \left(\ddot{\mathbf{r}}^i - \mathbf{a}^i\right) \cdot \left(\dot{\mathbf{r}}^i - \mathbf{w}_{ie}^i \times \mathbf{r}^i\right). \tag{2.39}$$

In *e*-frame, multiplying $\mathbf{r}^e$ to (2.36), I have

$$\ddot{\mathbf{r}}^e \cdot \dot{\mathbf{r}}^e = \nabla^e V^E \cdot \dot{\mathbf{r}}^e + \mathbf{a}^e \cdot \dot{\mathbf{r}}^e - 2\left(\mathbf{w}_{ie}^e \times \dot{\mathbf{r}}^e\right) \cdot \dot{\mathbf{r}}^e - \left[\mathbf{w}_{ie}^e \times \left(\mathbf{w}_{ie}^e \times \mathbf{r}^e\right)\right] \cdot \dot{\mathbf{r}}^e,$$

where the third term on the right-hand side is zero. Substituting the above equation into (2.38) by considering the second assumption, I can rewrite equation (2.38), the time-derivative of static geopotential in *e*-frame, as

$$\frac{dV^E}{dt} = \left[\ddot{\mathbf{r}}^e - \mathbf{a}^e + \dot{\mathbf{w}}_{ie}^e \times \left(\dot{\mathbf{w}}_{ie}^e \times \mathbf{r}^e\right)\right] \cdot \dot{\mathbf{r}}^e. \tag{2.40}$$

Clearly (2.39) and (2.40) are *equivalent* since they are both derived from the same equations and the same two assumptions; that is, one can easily show

$$\frac{dV^E}{dt} = \left(\ddot{\mathbf{r}}^i - \mathbf{a}^i\right)\cdot\left(\dot{\mathbf{r}}^i - \mathbf{w}_{ie}^i \times \mathbf{r}^i\right) = \left[\ddot{\mathbf{r}}^e - \mathbf{a}^e + \mathbf{w}_{ie}^e \times \left(\mathbf{w}_{ie}^e \times \mathbf{r}^e\right)\right]\cdot \dot{\mathbf{r}}^e = \left(\nabla^e V^E\right)^T \cdot \dot{\mathbf{r}}^e .$$

2.3.1.3 Energy Equation in Both Frames

The energy equation in *i*-frame is obtained by integrating (2.39) with respect to time as

$$V^E = \int_{t_0}^{t} \left(\ddot{\mathbf{r}}^i - \mathbf{a}^i\right)\cdot\left(\dot{\mathbf{r}}^i - \mathbf{w}_{ie}^i \times \mathbf{r}^i\right) dt .$$

Because of $\dot{\mathbf{w}}_{ie}^i = 0$, except that $\mathbf{a}^i$ term can't be explicitly integrated, all the other terms can be easily integrated as

$$\int_{t_0}^{t} \ddot{\mathbf{r}}^i \cdot \dot{\mathbf{r}}^i \, dt = \frac{1}{2}\left|\dot{\mathbf{r}}^i\right|^2 - E^{1i}$$

$$\int_{t_0}^{t} \ddot{\mathbf{r}}^i \cdot \left(\mathbf{w}_{ie}^i \times \mathbf{r}^i\right) dt = \mathbf{w}_{ie}^i \cdot \left(\mathbf{r}^i \times \dot{\mathbf{r}}^i\right) - E^{2i}$$

I arrive at the formulation of energy equation in *i*-frame as follows:

$$V^E = \frac{1}{2}\left|\dot{\mathbf{r}}^i\right|^2 - \mathbf{w}_{ie}^i \cdot \left(\mathbf{r}^i \times \dot{\mathbf{r}}^i\right) - \int_{t_0}^{t} \mathbf{a}^i \cdot \left(\dot{\mathbf{r}}^i - \mathbf{w}_{ie}^i \times \mathbf{r}^i\right) dt - E^{0i} . \tag{2.41}$$

Similarly, the energy equation in *e*-frame is obtained by integrating (2.40) with respect to time as

$$V^E = \int_{t_0}^{t} \left[\ddot{\mathbf{r}}^e - \mathbf{a}^e + \dot{\mathbf{w}}_{ie}^e \times \left(\dot{\mathbf{w}}_{ie}^e \times \mathbf{r}^e\right)\right]\cdot \dot{\mathbf{r}}^e \, dt .$$

Because of $\dot{\mathbf{w}}_{ie}^e = 0$ as well, only $\mathbf{a}^e$ term can't be explicitly integrated as the same, and other terms can be easily integrated as

$$\int_{t_0}^{t} \ddot{\mathbf{r}}^e \cdot \dot{\mathbf{r}}^e \, dt = \frac{1}{2}\left|\dot{\mathbf{r}}^e\right|^2 - E^{1e}$$

$$\int_{t_0}^{t}\left[\mathbf{w}_{ie}^{e}\times\left(\mathbf{w}_{ie}^{e}\times\mathbf{r}^{e}\right)\right]\cdot\dot{\mathbf{r}}^{e}\,dt=\int_{t_0}^{t}\left[\left(\mathbf{w}_{ie}^{e}\cdot\mathbf{r}^{e}\right)\left(\mathbf{w}_{ie}^{e}\cdot\dot{\mathbf{r}}^{e}\right)-\left(\mathbf{w}_{ie}^{e}\cdot\mathbf{w}_{ie}^{e}\right)\left(\mathbf{r}^{e}\cdot\dot{\vec{r}}^{e}\right)\right]dt$$

$$=\frac{1}{2}\left(\mathbf{w}_{ie}^{e}\cdot\mathbf{r}^{e}\right)\left(\mathbf{w}_{ie}^{e}\cdot\mathbf{r}^{e}\right)-\frac{1}{2}\left(\mathbf{w}_{ie}^{e}\cdot\mathbf{w}_{ie}^{e}\right)\left(\mathbf{r}^{e}\cdot\mathbf{r}^{e}\right)-E^{2e}$$

$$=-\frac{1}{2}\left|\mathbf{w}_{ie}^{e}\times\mathbf{r}^{e}\right|^{2}-E^{2e}$$

where the cross product and dot product are related by $\mathbf{a}\times\left(\mathbf{b}\times\mathbf{c}\right)=\mathbf{b}\left(\mathbf{a}\cdot\mathbf{c}\right)-\mathbf{c}\left(\mathbf{a}\cdot\mathbf{b}\right)$ and $\left|\mathbf{a}\times\mathbf{b}\right|^{2}=\left|\mathbf{a}\right|^{2}\left|\mathbf{b}\right|^{2}-\left(\mathbf{a}\cdot\mathbf{b}\right)^{2}$. And finally I arrive at the formulation of energy equation in *e*-frame as follows:

$$V^{E}=\frac{1}{2}\left|\dot{\mathbf{r}}^{e}\right|^{2}-\frac{1}{2}\left|\dot{\mathbf{w}}_{ie}^{e}\times\mathbf{r}^{e}\right|^{2}-\int_{t_0}^{t}\mathbf{a}^{e}\cdot\dot{\mathbf{r}}^{e}\,dt-E^{0e}\,. \tag{2.42}$$

Again, energy equation (2.41) and (2.42) are equivalent just as (2.39) and (2.40). Thus, as it should be, there is no difference no matter in which frame the energy equation is used for the energy balance method [*Zeng et al.*, 2015]. In this study, I adopt the equation in inertial frame (2.41), i.e., equation (2.31), for the GRACE data processing, since most of the measurements are given in that frame.

Finally, it should be noticed that in both frames $\int_{t_0}^{t}\mathbf{a}\cdot\dot{\mathbf{r}}\,dt\neq V^{R}+\int_{t_0}^{t}\mathbf{f}\cdot\dot{\mathbf{r}}\,dt$, since residual geopotential V^R (such as tides) is not static in either frames. I mention that some previous studies have already derived the identical [e.g. *Han*, 2003; *Gerlach et al.*, 2003; *Wang et al.*, 2012] or similar [e.g. *Badura et al.*, 2006; *Jäggi et al.*, 2008] formulation.

2.3.2 Numerical Assessment of the Formulation

2.3.2.1 Computation of angular velocity vectors in both frames

Both energy equations in *i*-frame and *e*-frame need the computation of angular velocity projected into each frame, i.e., coordinates vectors $\mathbf{w}_{ie}^{i}$ and $\mathbf{w}_{ie}^{e}$. There are at least two approaches to compute this vector. The first approach needs an intermediate frame in which the angular velocity only has non-zero values along z-axis. In this intermediate frame (called *m*-frame here), the z-axis is always in the same direction of instantaneous rotation axis. The angular velocity vector in *m*-frame can be written as $\mathbf{w}_{ie}^{m}=\begin{pmatrix}0 & 0 & \Omega\end{pmatrix}^{T}$, where Ω is the instantaneous rotation rate of *e*-frame with respect to *i*-frame. Therefore one could transform this coordinate vector from *m*-frame to *i*-frame and *e*-frame by

$$\begin{aligned} \mathbf{w}_{ie}^{i} &= C_{m}^{i}\mathbf{w}_{ie}^{m} \\ \mathbf{w}_{ie}^{e} &= C_{m}^{e}\mathbf{w}_{ie}^{m} \end{aligned}. \tag{2.43}$$

In practice, *i*-frame is realized by ICRF and *e*-frame is realized by ITRF. However, the intermediate frame is not explicitly provided. The best approximation of intermediate frame is the frame along the so-called Celestial Intermediate Pole (CIP). If the difference between the CIP and the instantaneous rotation axis is neglected, the transformation between *m*-frame and *i*-frame can be approximated by the procession and nutation matrix, and the transformation between *m*-frame and *e*-frame can be approximated by polar motion matrix.

The second approach is to use the skew-symmetric matrix from section 2.3.2, because the three non-zero entries for each skew-symmetric matrix represent the three coordinates of angular velocity vector in each frame. Considering the relationship of $\dot{C}_{e}^{i} = \Omega_{ie}^{i}C_{e}^{i} = C_{e}^{i}\Omega_{ie}^{e} = \left[\mathbf{w}_{ie}^{i}\times\right]C_{e}^{i} = C_{e}^{i}\left[\mathbf{w}_{ie}^{e}\times\right]$, one can easily derived the skew-symmetric matrix in both frames as

$$\begin{aligned} \Omega_{ie}^{i} &= \dot{C}_{e}^{i}C_{i}^{e} \\ \Omega_{ie}^{e} &= C_{i}^{e}\dot{C}_{e}^{i} \end{aligned}, \tag{2.44}$$

Compared to the first approach, this approach requires the time derivatives of the transformation matrix between the two frames. But such time derivatives are either not provided in the IERS standard. So in practice, numerical differentiation can be used to compute the time derivative, or if one has the knowledge of the rates of all the Earth orientation parameters (Earth rotation, procession, nutation and polar motion), the time derivative can also be explicitly derived.

In this study, I adopt the first approach to compute the angular velocity vectors in inertial frame.

2.3.2.2 Contribution of the Additional Terms From the Formulation

Comparing the formulation (2.31) and the original formulation (2.30), it is obvious that there are several terms missing from the original formulation, such as the term due to the angular velocity vector, and terms due to the time-variable residual geopotential, which has been neglected by the previous studies as an approximation. This approximation can be expressed as an additional term δV^{E} by subtracting equation (2.30) from (2.31) as

$$
\begin{aligned}
\delta V^E &= \left(\frac{1}{2}|\dot{\mathbf{r}}|^2 - \mathbf{w}\cdot(\dot{\mathbf{r}}\times\dot{\mathbf{r}}) - \int_{t_0}^{t} \mathbf{a}\cdot(\dot{\mathbf{r}} - \mathbf{w}\times\mathbf{r})\,dt - E^0 \right) \\
&\quad - \left(\frac{1}{2}|\dot{\mathbf{r}}|^2 - \omega(x\dot{y} - y\dot{x}) - \int_{t_0}^{t} \mathbf{f}\cdot\dot{\mathbf{r}}\,dt - V^R - E^0 \right) \\
&= \left[\omega(x\dot{y} - y\dot{x}) - \mathbf{w}\cdot(\dot{\mathbf{r}}\times\dot{\mathbf{r}}) \right] + \int_{t_0}^{t} \mathbf{a}\cdot(\mathbf{w}\times\mathbf{r})\,dt + \left[V^R - \int_{t_0}^{t} \nabla V^R\cdot\dot{\mathbf{r}}\,dt \right] \qquad ,(2.45) \\
&= \left[\omega(x\dot{y} - y\dot{x}) - \mathbf{w}\cdot(\dot{\mathbf{r}}\times\dot{\mathbf{r}}) \right] + \int_{t_0}^{t} \mathbf{f}\cdot(\mathbf{w}\times\mathbf{r})\,dt + \int_{t_0}^{t} \nabla V^R\cdot(\mathbf{w}\times\mathbf{r})\,dt + \left[\int_{t_0}^{t} \frac{\partial V^R}{\partial t}\,dt \right] \\
&= V_W + V_{WF} + V_{WR} + V_{TR}
\end{aligned}
$$

where V_W is denoted as the first term, representing the approximation due to the difference between the inertial frame where the orbit data are computed and the intermediate inertial frame, i.e., CIP frame, where angular velocity vector has non-zero values only along z-axis, V_{WF} is denoted as the second term, representing the approximation caused by both angular velocity vector and the non-conservative force, V_{WR} is denoted as the third term, representing the similar approximation but caused by the time-variable residual geopotential, and V_{TR} is denoted as the fourth term, which can be written in the form of 'potential rotation' term, representing effect caused by the rotation of the time-variable residual geopotential in inertial frame.

I conduct a simulation to calculate and assess each term, and evaluate if the approximation in the original formulation is negligible or not. Since the simulation is for GRACE constellation, all the potential terms in equation (2.45) are regarded as the potential differences. First, I integrate the equations of motion of two satellites to simulate the orbits, including positions and velocities, for a pair of satellites in GRACE configuration, and then use the satellite positions and velocities to compute each additional term according to equation (2.45), with the method of computing angular velocity vector according to equation (2.43). Computation is done for a day arbitrarily chosen (May 01, 2006). The dynamic model from Section 2.2.3 is also used here for the computation of residual geopotential terms. In addition, I decompose the residual geopotential terms (both V_{WR} and V_{TR}) into 4 parts, which are N-body perturbation, i.e., tide generating potential, solid Earth tide, ocean tides, and other effects. The last part includes mainly AOD1B, as well as pole tide, atmosphere tide and relativistic effect.

The time series for one day of all the ten terms (1 V_W term, 1 V_{WF} term, 4 V_{WR} terms, and 4 V_{TR} terms) from equation (2.45) are shown in Figure 2.2. Here the orbital revolution period, i.e., 5400 seconds, is used as the time scale for a better understanding of the main frequency of these terms. The unit of the y-axis is m^2/s^2, in terms of potential difference. It is worth mentioning that the typical magnitude of the signal from time-variable gravity field is usually at the level from 0.001 m^2/s^2 ~ 0.01 m^2/s^2 in terms of geopotential differences (see next section on the resulting geopotential difference estimates using real data). Therefore, any neglected terms larger than this threshold would significantly

contaminate the estimation of geopotential differences and the corresponding recovery of the gravity field.

It can be clearly seen that, among all the 10 terms the largest one is V_{W}, as shown in Figure 2.2 (a), the peak-to-peak amplitude of which is about 0.1 m^2/s^2. The dominant frequency of this term is 2-cycle-per-revolution (CPR), caused by the displacement from the poles of the two frames (orbit inertial frame and CIP frame), which means that neglecting of this term by using the original energy equation would certainly introduce 2CPR errors to the estimation of geopotential difference. Of course, Figure 2.2 (a) only shows this term for one particular day and the amplitude of this approximation actually could greatly vary depending on the time. For example, in practice the orbit data are normally computed in the frame of orbit frame J2000.0, but for case of Figure 2.2 (a) the CIP has shifted from J2000.0 to the epoch of the day, i.e. May 01, 2006. More than six years' precession would cause very large displacement between the two poles, and such displacement between the pole on J2000.0 and the CIP on May 01, 2006 would cause very large 2 CPR errors for V_{W} term. That means the amplitude of this term would become larger when the epoch of the measurement is far away from J2000.0 and smaller when the epoch is close to J2000.0. On the other hand, this approximation can be also avoided if one could properly handle the orbit data. For instant, if one would first transform the orbit data from J2000.0 to CIP frame and then apply the energy equation, then the V_{W} term would be zero.

Figure 2.2 (b) shows that the V_{WF} term has the peak-to-peak amplitude about 0.01 m^2/s^2, which is also non-negligible compared to the magnitude of the time-variable signal. The frequency of this term should be modulated by the non-conservative force **f**, which is measured by the on-board accelerometer and star camera. The typical frequency of non-conservative force should be 1CPR plus measurement noise, and so the integral of non-conservative force should yield a term with frequency of 1CPR plus random walks, just like the time series shown in Figure 2.2 (b).

The 4 V_{WR} terms, Figure 2.2 (c)~(f), are similar to V_{WF} term, except that they are driven by the corresponding tidal potential. The N-body perturbation and solid Earth tide both have a dominated frequency of 2CPR because they mainly produce time dependent change in degree 2 coefficients and also GRACE satellites are flying in a near circular polar orbit. Therefore, the first two time series, Figure 2.2 (c) and (d), from these two effects have a similar 2 CPR frequency. The third term (Figure 2.2 (e)) from ocean tides and the forth term (Figure 2.2 (f)) mainly from AOD1B show time series with more complicated frequencies because ocean tides and AOD1B produce time dependent changes in much higher degrees and orders besides degree 2. Then about the amplitude for each term, the three terms, from N-body, solid Earth tide and ocean tides, are all above the signal level of 0.001 m^2/s^2 ~ 0.01 m^2/s^2, and thus can not be ignored.

The 4 V_{TR} terms, which can be also written as $\int_{t_0}^{t} \frac{\partial V^R}{\partial t} dt$, represent the integral of the partial derivative of residual geopotential with respect to time in inertial frame, according

to equation (2.37). Therefore, the magnitude of these terms can be simply estimated by evaluating the partial time derivative of the corresponding potential in inertial frame. For example, the magnitude caused by the partial time derivative of the potential from N-body and solid Earth tide should be much smaller than the potential itself, but for ocean tides and AOD1B, the corresponding magnitude of the time derivative should be almost at the same level of the potential. The former is because the time variation of N-body effect and solid Earth tide is mainly driven by the motion of celestial bodies, such as the sun and the Moon, which is quite slow in inertial frame, and thus yield smaller magnitude as shown in Figure 2.2 (g) and (h). But for ocean tides and AOD1B, the time variation is caused by both the higher degree and order of those tides and earth rotation, and therefore ocean tides produce larger time variation than the N-body and solid Earth tide as shown in Figure 2.2 (i) and AOD1B produce almost the same level as N-body as shown in Figure 2.2 (j). Comparing to the signal threshold of 0.001 m^2/s^2 ~ 0.01 m^2/s^2, I conclude that at least the effect from ocean tides, with the peak-to-peak magnitude of about 0.005 m^2/s^2, should not be neglected from the original formulation.

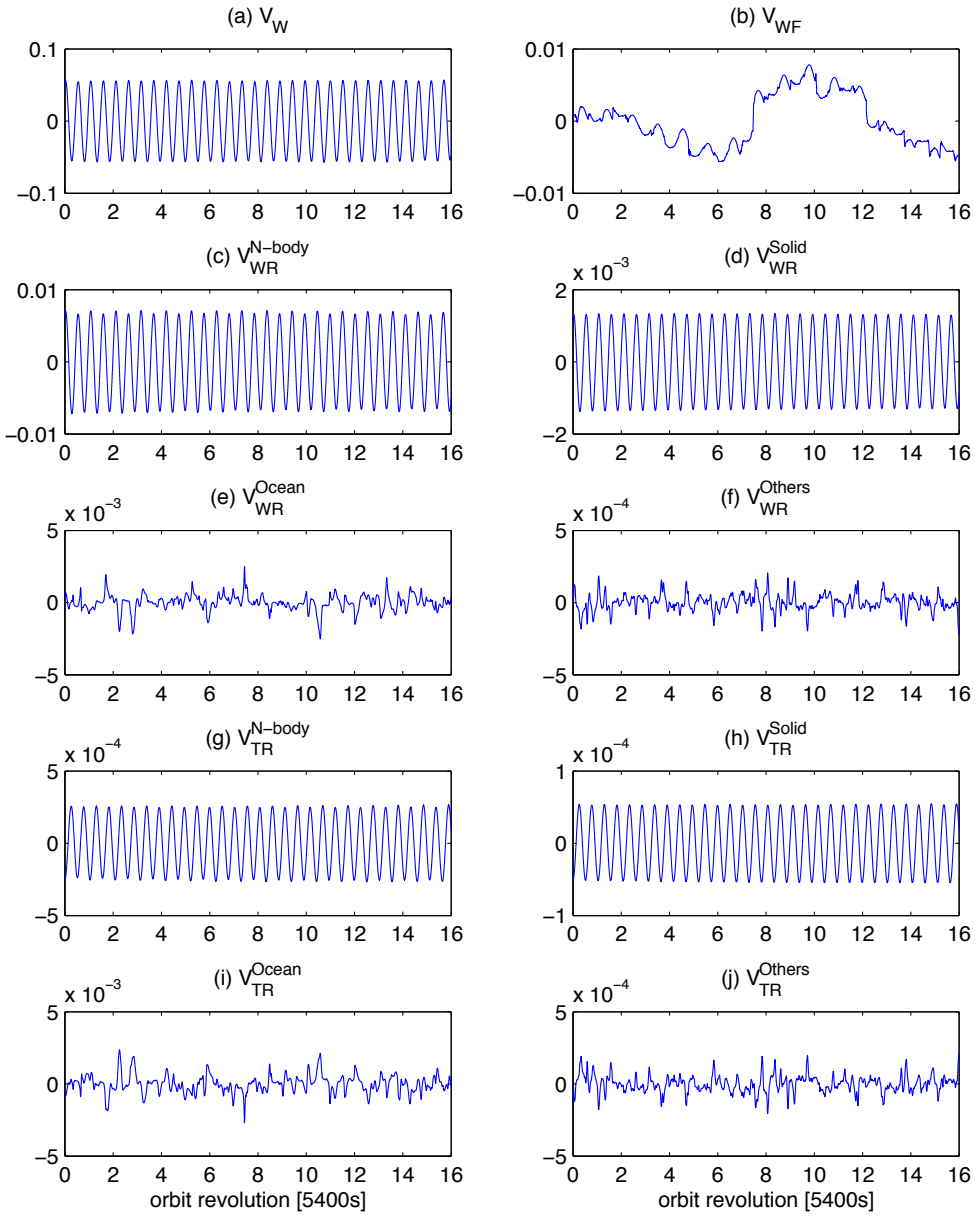


Figure 2.2 Comparison of the contribution of different terms in equation (2.45). (a) V_W, the approximation due to the difference between the J2000.0 inertial frame where the orbit data are computed and the CIP inertial frame where angular velocity vector has non-zero values only along z-axis. (b) V_{WF}, the approximation caused by both angular velocity vector and the non-conservative force. (c)~(f) V_{WR}, similar to V_{WF} term, but caused by N-body perturbation, solid Earth tide, ocean tides and others (including mainly AOD1B, also pole tide, atmosphere tide and relativistic effect), respectively. (g)~(j) V_{TR}, 'potential rotation term' caused by the rotation of the time-variable residual geopotential in inertial frame, for N-body, solid Earth tide, ocean tides, and others (including mainly AOD1B, also pole tide, atmosphere tide and relativistic effect), respectively.

2.3.2.3 The Accuracy of the Formulation

Although the formulation (2.31) has made improvement by including all the additional terms neglected by the original formulation (2.30), according to the deviation from Section 2.3.2, two approximations are still inevitable and have to be assumed during the deviation of equation (2.31). One is to assume V_E to be static during the integral limits from t_0 to t, which is consistent with the GRACE convention, i.e., estimating a mean gravity field during a certain time interval, so it will not introduce any mathematical error by using the formulation. The other one is to assume the rates of Earth's angular velocity vector is zero, i.e., $\dot{\mathbf{w}} = 0$, which indeed would cause errors but negligible compared to measurement noise level of GRACE. In this subsection, a closed-loop simulation is used to assess the accuracy of formulation (2.31) by evaluating this approximation. For comparison the original formulation (2.30) is also assessed but with the assumption that the largest approximation caused by the V_W term has been avoided, i.e., the orbit data have been appropriately transformed from the J2000.0 frame to the CIP frame.

First purely dynamic orbits for two satellites are simulated as the input to the energy equation. As mentioned in Section 2.2.1, using a purely dynamic orbit as the input data of energy equation should reduce the estimates to the *a priori* gravity field. If the residual geopotential difference observables are defined as: $\Delta V_{12}^E = V_{12}^E - V_{12}^{E\_aprior}$, i.e., the difference between the estimated values using energy equation and the predicted values using the *a priori* gravity model, then theoretically the residual should be reduced to zero, if the formulation of energy equation is absolutely accurate. Therefore the nonzero residual would reveal the approximate errors caused by using each formulation of energy equation.

The residuals based on equation (2.30) and (2.31) are presented in Figure 2.3, for the integral limit of one day. Figure 2.3 (a) shows the residuals based on both equations, and Figure 2.3 (b) zooms in on the residual based on formulation (2.31) only. From Figure 2.3 (a), one can see that the one based on formulation (2.31) is reduced to almost zero, but the one based on original formulation (2.30) contains a relatively large nonzero residual with a dominated 2CPR (2-cycle-per-revolution) error, which is mainly caused by neglecting all the terms from Figure 2.2 (b) to (j). The magnitude of the error is just the magnitude of the summation of all the time series from Figure 2.2 (b) to (j), which is about 0.02 m^2/s^2 from peak to peak. The errors caused by the approximation would definitely overcome the signal level from the time-variable gravity field, and would further corrupt the geopotential difference estimates as well as the recovery of gravity field.

Although the dominated frequency of the errors is close to 2CPR, which seems to be able to be removed by an additional 2CPR parameter [*Han et al*., 2006; *Tangdamrongsub et al*., 2012] or even more parameters [*Ramillien et al*., 2011], but actually the errors contain much more high-frequency constituents (e.g., from ocean tides) that the empirical

parameters cannot fully absorb. In addition, because of the polar-orbit configuration, the 2CPR empirical parameters would highly correlate with the degree 2 gravity coefficients, and the same for other CPR empirical parameters. Therefore, using too many CPR empirical parameters would surely contaminate gravity signal, especially the zonal geopotential coefficients. In addition, the studies based on traditional orbit dynamic approach also indicate that the empirical parameterization should be no more than 1 CPR [*Tapley et al.*, 2004a] or with even less empirical parameters [*Luthcke et al.*, 2006], besides bias and trend, which is for the purpose to mitigate the systematic error of range-rate data, as well as to better retain the time-variable geopotential signal. Therefore, I conclude the using large number of empirical parameters in the previous studies based on energy balance method might be actually caused by the systematic error from the original formulation of energy equation, and in order to fully exploit the precision of GRACE data, it is requisite to choose equation (2.31) as the practical formulation of energy equation.

A zoomed view of the error from the new formulation is presented in Figure 2.3 (b), even though it appears to be negligible as the red line shown in Figure 2.3 (a). One can see that the error is mainly composed of a liner component and a 2CPR component, which is caused by both the assumption of $\dot{\mathbf{w}} = 0$ and the approximate realization of intermediate frame using CIP frame. The linear component will be removed using a trend parameter so it will not bring any additional errors. The 2CPR component does introduce systematic error, but only with the peak-to-peak amplitude of the order less than 5×10^{-5} m^2/s^2, which is definitely negligible for current GRACE measurement accuracy and probably also for GRACE follow-on measurement accuracy in the future [*Loomis et al.*, 2012].

The detailed numerical evaluations can be found in *Guo et al.* [2015] on the discussion in the inertial frame and *Zeng et al.* [2015] on the discussion in earth-fixed frame formulations, respectively.

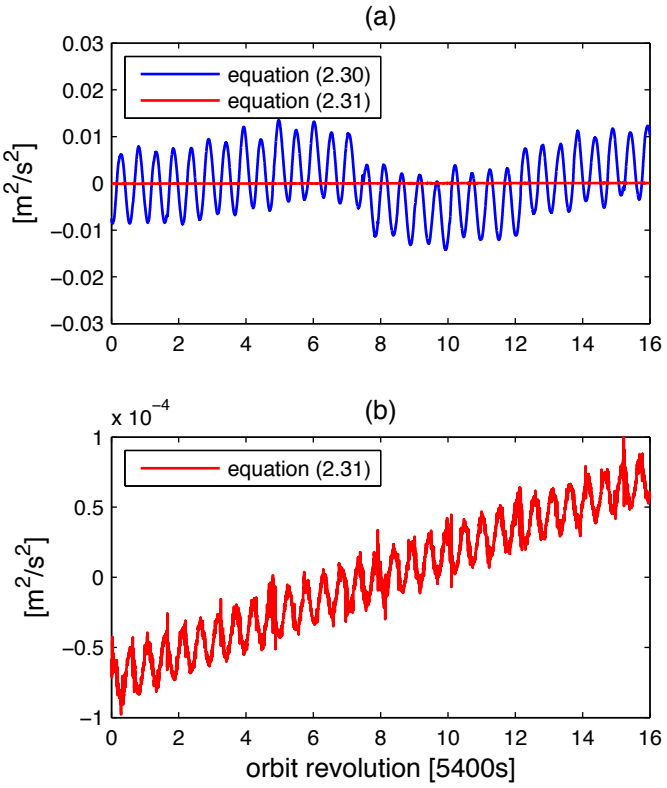


Figure 2.3 Residual geopotential differences from a closed-loop simulation based on equation (2.30) and equation (2.31). Simulated dynamic orbit data are only used as the input, so the residuals compared to a priori gravity field should be reduced to zero if the formulation is absolutely accurate. (a) Residual geopotential differences based on equation (2.30) and (2,31). The blue line represents the residuals based on equation (2.30) and the red line represents the residuals based on equation (2.31). (b) Zoomed view of (a), but for only residual geopotential difference based on equation (2.31).

Chapter 3 GRACE Time-variable Gravity Solutions and Applications

3.1 Geopotential Difference Estimates

As the exclusive approach that can directly link each KBR measurement to a geophysical quantity, i.e., geopotential difference between the two positions of the twin satellites, energy balance approach could provide a unique aspect to extend our conventional knowledge about both data processing and results interpretation of GRACE. The geopotential difference estimates are able to sense the gravity information, without losing any high-frequency resolution, since each estimate is computed straight from range-rate measurement for each epoch. Because of that, the geopotential difference estimates could not only be used for both global and regional gravity recovery but also be regarded as an in-situ gravity representation without downward continuation [*Han et al.*, 2006]. Therefore, an accurate estimation of geopotential difference observables is the key issue for energy balance approach and is the most critical step for the subsequent temporal gravity inversion.

In order to estimate geopotential differences, the input data from GRACE L1B products [*Case et al.*, 2010; *Wu et al.*, 2006] are processed based on the methodology and full background model described in last chapter. The procedures can be summarized as follows:

1. GNV1B orbit data that are used to reconstruct the purely dynamic orbit and estimate the daily accelerometer calibration parameters, with respect to the full background model.

2. Range-rate data from KBR1B product are used to correct the relative velocity components of the reconstructed orbit via the alignment equation (2.4), i.e., from reconstructed relative velocity vector $\dot{\mathbf{r}}_{12}$ to new 'pitch-free' relative velocity vector $\dot{\mathbf{r}}_{12}^{\rho}$.

3. Energy equation is applied according to the formulation as equation (2.31) to compute the raw geopotential difference observables. The background models are used again for computing all the correction terms, and the accelerometer calibration parameters obtained from orbit reconstruction are also used.

4. Finally, since the raw geopotential differences would inevitably contain systematic error inherited from the range-rate data, a number of empirical parameters are estimated and removed from the raw geopotential difference observables. Again, because of the improved methodologies, the empirical parameters only contain bias, rate and 1 CPR

parameters for every orbital revolution, which is consistent with other approaches of GRACE data processing commonly used in other studies.

According to the procedures described above, energy balance method can be regarded as a kind of 'two-step' approach, i.e., adjustment of the reference orbits first (Step 1) and estimation of gravity (geopotential difference) using the fixed reference orbits next (Step 4), which is different from the term 'two-step' used in traditional dynamic method, where it means adjustment of GPS orbits in the first step and estimation of both GRACE orbits and time-variable gravity in the second step with the GPS orbits fixed [*Dahle et al.*, 2012]. Two calibrations of the measurements are needed in both Step 1 and Step 4, for accelerometer and range-rate data. In the following subsections, the two calibrations will be first discussed and then the result of the estimated geopotential differences will be presented.

3.1.1 Calibration of Accelerometers

The on-board accelerometers are designed to accurately measure the non-conservative force, including atmosphere drag, solar radiation pressure and spacecraft propulsion. Each satellite is equipped with a SuperSTAR accelerometer at the center of mass of the satellite. The accelerometer measures the non-conservative force in three orthogonal directions of the Science Reference Frame (SRF), which can be transformed into inertial frame using the quaternions from the SCA1B product measured by the Star Camera Assembly. Approximately, X_{SRF} is along roll axis in the anti-flight and in-flight directions for the leading and trailing satellites, respectively, Z_{SRF} is along the yaw axis and points to nadir, and Y_{SRF} is along pitch axis and forms a right-handed triad with X_{SRF} and Z_{SRF} [*Case et al.*, 2010].

Despite the high precision of about 10^{-10} m/s^2 within the bandwidth of 2×10^{-4}–0.1 Hz [*Kang et al.*, 2006], the raw accelerometer data directly from ACC1B product suffer from large uncertainty at lower bandwidth ($<2\times10^{-4}$ Hz), and therefore need to be calibrated first. A widely used method is to treat those low frequency errors as a kind of systematic error and calibrate them by estimating a group of empirical parameters, such as bias and/or scale parameters. The bias calibration usually includes not only the offset, but also the linear and even quadratic trend. The scale calibration is mainly to correct the additional 1CPR error, similar to all the data measured by the onboard payloads. Currently, most of the data processing centers adopt the strategies to calibration both bias and scale parameter daily or sub-daily. However, *Van Helleputte et al.* [2009] found that strong correlation may exist between scale and bias, and implied that the daily estimation of both bias and scale together may lead to unrealistic large variation of scale parameters compared to the non-conservative force model. Besides, unlike the 'one-step' dynamic method used by other processing centers, energy balance method is a 'two-step' method, which means the orbit has to be adjusted with the accelerometer parameters first (Step 1) and then next (Step 4) the geopotential difference is estimated by using range-rate data. In that case, the daily scale parameters of accelerometer may be also correlated with the 1CPR parameter of range-rate. Therefore, in this study, a different strategy is adopted to

calibrate the accelerometer with the main purpose to suppress the unrealistic large day-to-day variation of scale parameters.

First, the *a priori* bias and scale parameters recommended by *Bettadpur* [2009] are adopted. Note that the *a priori* values are estimated based on analysis of data only between launch and March 31, 2009. Then after correcting the accelerometers using the *a priori* values, I estimate the daily bias (offset, linear and quadratic trend) and scale parameters between launch and end of 2013 using the orbit reconstruction method. The estimated scale parameters are shown in Figure 3.1. As indicated by *Van Helleputte et al.* [2009], very large variation can be found on all the directions of the estimated scale for both satellites, which may be unrealistic since there is no reason to believe the scales should be significantly different for different days [*Bettadpur*, 2009]. Another apparent feature is that flight direction (X_{SRF}) shows less uncertainty than the other two directions, which is consistent with the a-priori values of the uncertainty from *Bettadpur* [2009], i.e., 0.002 in X_{SRF} direction and 0.02 in the other two directions. It is not a surprise because along-track is the direction that would sense most of the non-conservative force compared to radial and cross-track directions. Therefore, here I only focus on the X_{SRF} direction and simply fix the scales in other two relative trivial directions using the *a priori* values.

From the time series of the X_{SRF} direction, it can be noticed that there seems to be an offset from 1.0 after 2010. In order to see more evidently, the yearly mean scales with the standard deviation are computed using all the estimated daily scale parameters. The yearly estimated scales of the X_{SRF} direction are shown in Figure 3.2. It can be seen clearly that the scale parameter jumps from approximate 1.0 to approximate 0.98 after 2010 with even less uncertainty compared to previous years. Therefore, the strategy in this study is to additionally adjust the X_{SRF} scale parameter to 0.98 after 2010, and keep 1.0 before 2010. It is worth mentioning that almost at the same time, i.e., at the end of 2010, the so-called 'battery management' started to operate for both satellites. Approximate every 5 months, the two satellites have to be switched off for about one month in order to reduce the load on the batteries. Therefore, I suspect that these operational events might be associated with the jump in the scale parameters of accelerometers.

In order to show this adjustment is necessary, the post-fit residuals of the reconstructed orbit are computed based on difference accelerometers calibration strategies. Since here the input data are the GNV1B orbit data, the post-fit residuals also represent the orbit difference between GNV1B and the reconstructed orbit. GNV1B orbit is a kind of reduced dynamic orbit, and therefore it is relatively free of accelerometer information. The results are shown in Figure 3.3 in terms of the daily root mean square (RMS) of the orbit difference. Here the same method is adopted to calibrate bias parameters as GFZ RL05 standard [*Dahle et al.*, 2012], i.e., estimate daily bias with offset, linear and quadratic trend. In the top row of Figure 3.3, the scale parameter is not estimated and just the *a priori* value is used. In that case, the RMS of the orbit difference rapidly increases after 2010. For comparison, in the middle row of Figure 3.3, the scale parameters are

estimated for each day, so the large jumps after 2010 are suppressed. Eventually, in the bottom row of Figure 3.3, the scale parameters are not daily estimated but the yearly mean scale parameters are used, i.e., adjust the X_{SRF} scale parameter to 0.98 after 2010 and keep 1.0 before 2010. As a result, the method adopted in this study can efficiently reduce the large jumps of RMS after 2010, which is similar to the strategy of directly calibrating scales, but the gain is the unrealistic large variation of daily-estimated scale parameters can be avoided.

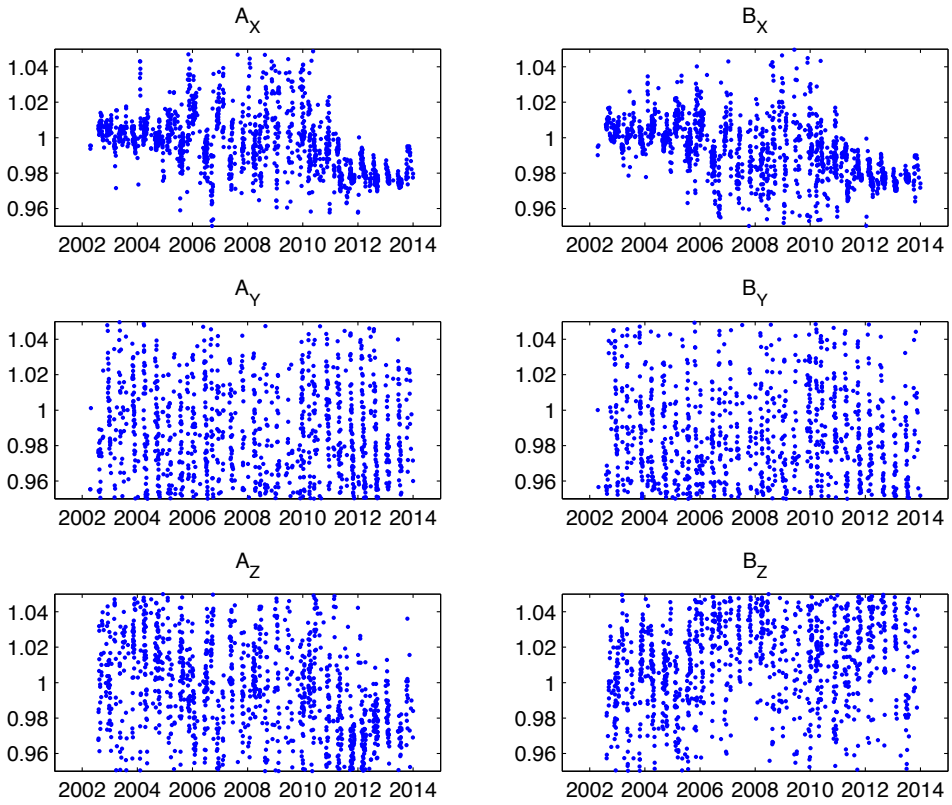


Figure 3.1 Daily estimated scale parameter of accelerometers for both satellites in three directions. All directions show large variation. It seems there exist an offset from 1.0 after 2010 in X_{SRF} direction for both satellites.

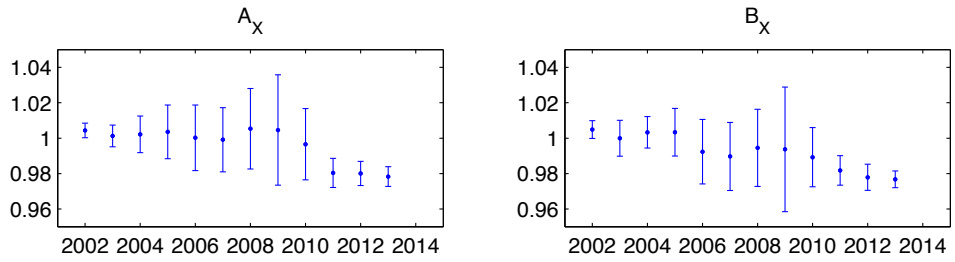


Figure 3.2 Yearly mean and error of the daily estimated scale parameters in X_{SRF} direction for both satellites. Scale parameter jumps from approximate 1.0 to approximate 0.98 after 2010 with less uncertainty.

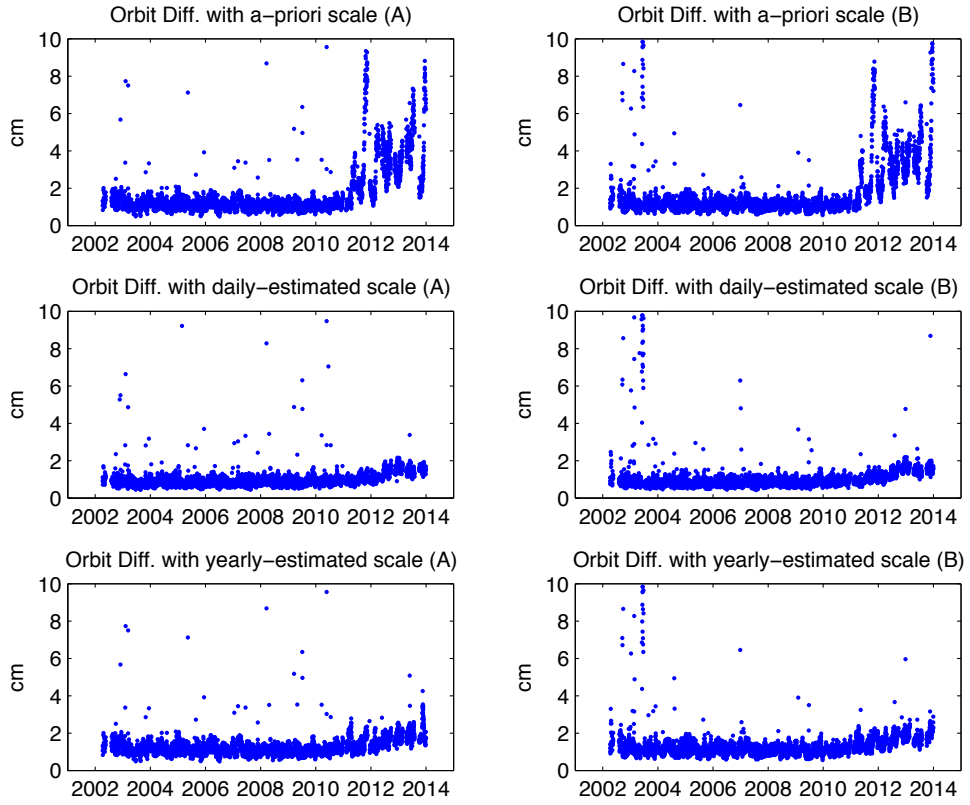


Figure 3.3 Root Mean Square (RMS) of the orbit difference between the input GNV1B orbit and output reconstructed dynamic orbit, using difference strategies to calibrate scale parameters of accelerometers. The orbit differences also represent the post-fit residuals. The top row: use a-priori value of the scale parameters. The Middle row: use daily-estimated scale parameters. The bottom row: use the yearly-mean scale parameters, i.e., i.e., adjust the X_{SRF} scale parameter to 0.98 after 2010 and keep 1.0 before 2010.

3.1.2 Calibration of Empirical Parameters

After the calibration of accelerometers as well as the reconstruction of the orbits, the range-rate measurements are used to align the orbit and then geopotential difference observables can be estimated using the energy equation. However, the direct estimated geopotential differences do not only contain the time-variable gravity signal, but also very large systematic error. An example of the raw geopotential differences is illustrated in Figure 3.4 (a) as a time series for the day of May 1, 2006, where referenced values from a static gravity field GIF48 [*Ries et al*., 2011] have been removed. It can be clearly seen from Figure 3.4 (a) that the time-variable gravity signals, i.e., the high frequency fluctuation, are almost invisible since they are overwhelmed by both an offset and very large 1CPR signals with 'butterfly' shape. The offset is trivial, which is caused by the difference of the integral constants between the two satellites, therefore it can be safely removed since gravity constant, i.e., degree 0 term, is usually not part of the GRACE solution. Similarly, the 1CPR signals are partly related to degree 1 term, which is routinely not estimated by GRACE either. However, the majority of these 1CPR signals, I conclude, are actually caused by two error, therefore, they must be treated as a kind of systematic error and have to be first removed using some empirical parameters before temporal gravity estimation. One error is the 1CPR systematic error directly inherited from KBR instrument, which should be small with relatively constant amplitude during a short period. The other error is caused by the imperfect background modeling used during orbit reconstruction. As we all know, orbit data obviously have a dominated frequency of 1CPR, and therefore any un-modeled or mis-modeled part in both reference fields and the accelerometers parameters would alias into same 1 CPR frequency during the orbit adjustment process. If such orbits were compared to more accurate data, such as the K-Band range-rate data, the residual of the time series would exhibit as 1CPR with 'butterfly' pattern for the whole the orbit arc as shown in Figure 3.4 (a). Therefore, the 1CPR part has to be removed as a systematic error since it is a kind of aliasing error and not directly from range-rate signals.

The systematic errors can be calibrated by also using a set of empirical parameters. There are at least two ways to estimate these parameters [*Zhao et al.*, 2011]. One is to estimate them simultaneously with gravity parameters. The other one is to estimate them directly from the raw residual time series before estimating the gravity field. Here the latter one is chosen in this study. One of reasons is that the purpose here is to first generate ready-to-use geopotential difference estimates before solving gravity field, so the processed geopotential difference estimates can directly be applied for other purposes, such as regional gravity analysis. But one should keep in mind that the either way would cause the correlation between the systematic error removed and the gravity signal recovered, which actually leads to a kind of signal-coherent adjustment. In the future, iteration may be considered to overcome this problem.

The adopted empirical parameters here include bias, trend and two 1CPR coefficients. The residuals after calibration of those empirical parameters are shown in Figure 3.4 (b). The figure shows the time series of the estimated geopotential difference residuals for the day of May 1, 2006. Again, the reference field is GIF48. The resulting geopotential differences is sampled every 5 seconds with the same sampling rate of range-rate data.

The root mean square (RMS) of the residuals is about 0.0024 m^2/s^2. If the small fluctuation in that time series can actually represent the geopotential change due to the time-variable gravity signal between May 1, 2006 and the reference epoch of GIF48, i.e., around 1 January 2007, it can be thus directly used for gravity field recovery.

Finally, we mention that low-pass filter can be another opinion for calibration of those systematic errors. The cutoff frequency should be set to a frequency lower than 2CPR in order to keep degree 2 signals. But the drawback is that edge effect is normally inevitable by using a low-pass filter, which leads to the possible data loss at each end of the orbit arc, i.e., the beginning and the end of each day. So similar to other GRACE data centers, empirical parameters method is still chosen over low-pass filter method in this study.

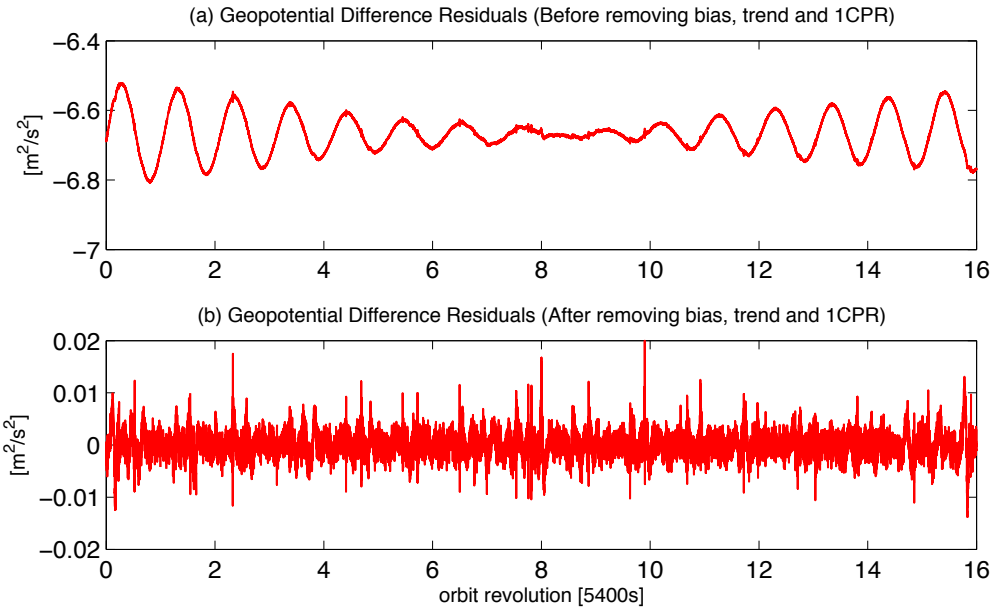


Figure 3.4 Geopotential difference estimates (with mean field removed) for the day of May 1, 2006. (a) Direct results by using alignment equation and energy equation without calibrating any empirical parameters. (b) Results after removing bias, trend and 1 CPR parameters.

3.1.3 Estimated Geopotential Differences

In order to carefully examine whether or not the resulting geopotential difference residuals can actually represent time-variable gravity signal, the geopotential difference estimates are compared to the predicted values from GRACE L2 product in Figure 3.5. Figure 3.5 (a) highlights the two ascending profiles of geopotential difference observables on a global map of the daily ground tracks of July 17, 2003. The coordinate of each estimate on the map is assigned to the middle point of two satellites and the color represents its value with a mean reference gravity field removed [*Ries et al.*, 2011]. In Figure 3.5 (b) and Figure 3.5 (c), the two profiles of estimates are respectively illustrated

with respect to latitude and compared to the predicted values using GRACE L2 solutions (from CSR RL05, GFZ RL05a, and JPL RL05, truncated to degree and order 60 with the same mean field removed) along the same profiles. First, it can be clearly seen that the geopotential differences estimates based on the approach in this study are similar to the predicted values from the three official L2 solutions. Of course, the three series of predicted values look smooth since they are simply computed from existing models with band-limited field up to degree and order 60 only, while the series of estimates is noisier because it is directly from range-rate measurements, which also implies that the new geopotential difference estimates could contain more detailed high-frequency information of the geopotential signal (as well as noise) directly sensed by the range-rate measurements.

As for each profile, Figure 3.5 (b) illustrates the profile approximately along 60°W longitude, mostly above the rough land area, and Figure 3.5 (c) is for the profile approximately along 140°W longitude, mostly above the flat ocean area. With the ascending of the satellite pair, the estimates in Figure 3.5 (b) obviously reveal the geopotential difference variation successively caused by West Antarctica, Amazon Basin, Hudson Bay and North Greenland, and the estimates in Figure 3.5 (c) mostly cover the Pacific Ocean. The surface gravity change can be directly inferred just from the geopotential difference profile. Because when the satellite pair passes by a gravity negative anomaly on ground, the geopotential differences (defined as the following satellite subtracting the leading satellite here) would exhibit increasing values first and decreasing values next, and vice versa. For example, in Figure 3.5 (b), from about 30°N to North Pole, a increase-decrease-increase fluctuation of geopotential difference can be observed, meaning that there should be a negative gravity anomaly followed by a positive anomaly compared to the epoch of the mean reference field (1 January 2007), which just corresponds the glacial isostatic adjustment (GIA) signal (negative) in the Hudson Bay, and Greenland ice sheet ablation signal (positive).

In Figure 3.5 (d), the time series for the day of May 1, 2006 (the same as Figure 3.4 (b)) is shown again but with the predicted time series from CSR L2 RL05 solution for comparison. Again, the time series of the estimates seems very close to the predicted values from CSR solution. Quantitatively speaking, the correlation coefficient of the time series between the estimates (after a 20-point moving average smoothing) and the prediction is about 0.91 for that certain day. As for all the estimates from 2003 to 2013, the average value of the daily correlation coefficients is over 0.9, which is much higher than correlations of 0.5~0.8 reported in previous study [e.g., *Han et al*., 2006].

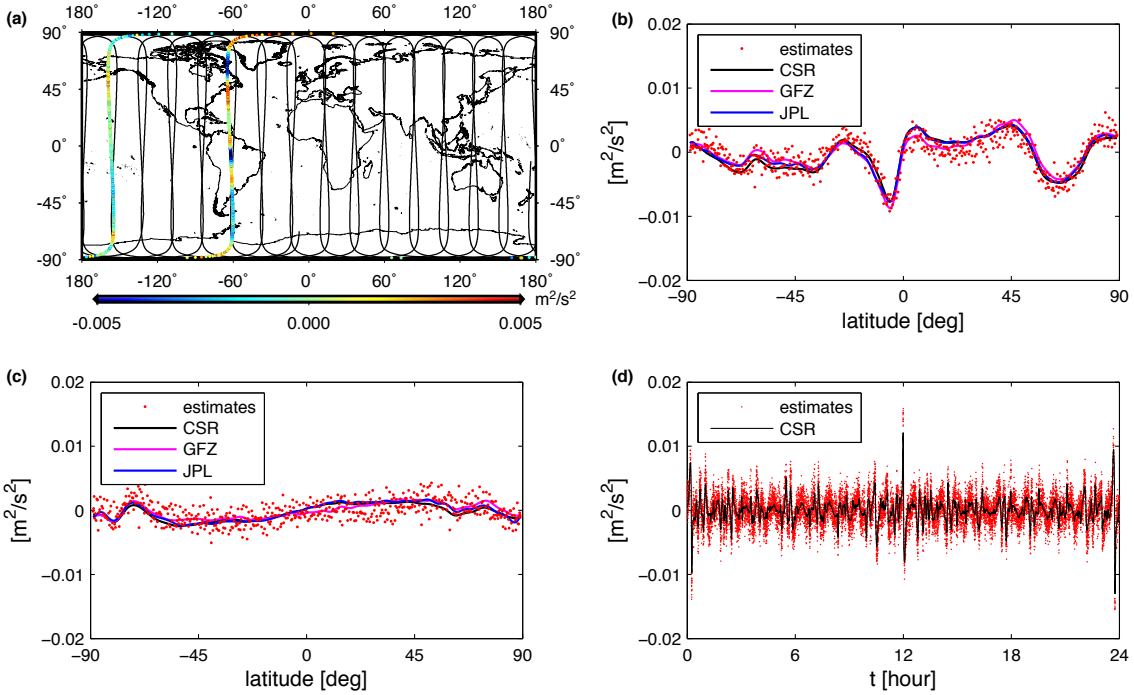


Figure 3.5 Geopotential difference estimates (with mean field removed). (a) Ground track of July 17, 2003 with two ascending profiles highlighted with color representing the values of the estimates. (b) Highlighted profile approximately along 60°W longitude mostly above the land area, with predicted values from GRACE L2 solutions (CSR RL05, GFZ RL05a, JPL RL05). (c) The same but for highlighted profile approximately along 140°W longitude mostly above the ocean area. (d) Time series of estimates for the day of May 1, 2006 with predicated values from CSR RL05 solution.

3.2 Monthly Global Gravity Solutions Using Geopotential Differences

As an in-situ observation type, geopotential differences can be directly applied to recover time-variable gravity for both regional signal and global signal. Previous researchers have already published several studies about the application of geopotential differences, but all of them are limited to regional gravity recovery only [e.g., *Han et al.*, 2005; *Schmidt et al.*, 2006, 2008; *Ramillien et al.*, 2011; *Tangdamrongsub et al.*, 2012]. The regional gravity recovery from geopotential differences, because of the non-uniqueness due to the downward continuation, requires additional constraints to stabilize the solution, such as regularization [*Han et al.*, 2005; *Tangdamrongsub et al.*, 2012] or various different basis functions [*Schmidt et al.*, 2006, 2008; *Simons et al.*, 2006], but still may be more easily and efficiently implemented compared to the traditional nonlinear dynamic orbit determination and parameter recovery method. However, the global solution from geopotential differences, though it is more straightforward because of the linear relationship between geopotential differences and Stokes coefficients, has not been successfully implemented before. Part of the reasons might be due to the systematic error due to the reference orbit and the approximation error in the previous energy equation. Because of the improvement from my energy balance method in this study, it becomes

possible to utilize geopotential differences to recover global gravity. In this section, the global gravity recovery will be studied using the improved geopotential difference estimates.

3.2.1 Inversion Method and Parallel Algorithm

Similar to the convention of official GRACE L2 product, a series of monthly mean solutions is solved for each calendar month, which means the gravity variation within a month has to be neglected and the recovered solution is assumed to be able to represent the mean gravity field of that month. The geopotential differences are accumulated during a month, which serve as an input for global gravity inversion. The output of the inversion is the set of Stokes coefficients, also similar to official GRACE L2 product.

The classic Gauss-Markov Model (GMM) is used to model this inversion problem straightforwardly. The relation of the geopotential difference V_{12}^E and Stokes coefficients ($\bar{C}_{nm}$ and $\bar{S}_{nm}$) can be expressed as

$$V_{12}^E = \frac{GM}{R}\sum_{n=2}^{n_{max}}\sum_{m=0}^{n}\left(\alpha_{nm}\bar{C}_{nm} + \beta_{nm}\bar{S}_{nm}\right), \tag{3.1}$$

where GM is the geocentric gravitational constant and R is Earth's mean spherical radius, n and m are degree and order, respectively, and n_{max} is the maximum degree. Here degree 0 and degree 1 coefficients are excluded, as GRACE range-rate measurements are insensitive to these parameters. The coefficients α_{nm} and β_{nm} are defined as

$$\begin{Bmatrix}\alpha_{nm}\\ \beta_{nm}\end{Bmatrix} = \left(\frac{R}{r_2}\right)^{n+1}\bar{P}_{nm}\left(\cos\theta_2\right)\begin{Bmatrix}\cos\left(m\lambda_2\right)\\ \sin\left(m\lambda_2\right)\end{Bmatrix} - \left(\frac{R}{r_1}\right)^{n+1}\bar{P}_{nm}\left(\cos\theta_1\right)\begin{Bmatrix}\cos\left(m\lambda_1\right)\\ \sin\left(m\lambda_1\right)\end{Bmatrix}, \tag{3.2}$$

where $\left(r_1,\theta_1,\lambda_1\right)$ and $\left(r_2,\theta_2,\lambda_2\right)$ are denoted as the spherical coordinates of the two satellites in Earth-fixed reference system, and $\bar{P}_{nm}$ is the fully normalized Legendre function. According to equation (3.1), the least squares principle to estimate the Stokes coefficients ($\bar{C}_{nm}$ and $\bar{S}_{nm}$) using data of geopotential difference V_{12}^E is

$$\sum_k\left\{\left(V_{12}^E\right)_k - \frac{GM}{R}\sum_{n=0}^{n_{\max}}\sum_{m=0}^{n}\left[\left(\alpha_{nm}\right)_k\bar{C}_{nm} + \left(\beta_{nm}\right)_k\bar{S}_{nm}\right]\right\}^2 = \min, \tag{3.3}$$

where the subscript k is used to represent data of the k-th observation including values of $\left(V_{12}^E\right)_k$, as well as $\left(\alpha_{nm}\right)_k$ and $\left(\beta_{nm}\right)_k$ computed using satellite coordinates according to equation (3.2). Based on the least squares principle, the solution of the unknowns ($\bar{C}_{nm}$

and $\bar{S}_{nm}$) can be easily solved from a large number of observations, and the corresponding formulation is the same as equation (2.8) and thus not reproduced here.

Here, the Cholesky decomposition is employed to invert the normal matrix, for the purpose to obtain both the estimates and the error covariance matrix. If the maximum degree and order is less than 60, the computation can be done using a normal workstation. If maximum degree and order were larger than 60, then the inversion would be more complexity, and the computation may be limited by both the float point operations (FLOPs) and memory storage. For example, assuming the maximum degree and order to be $n_{\max}$, the number of the total coefficients would be $M=(n_{\max}+1)\times(n_{\max}+1)$. So the FLOPs of the Cholesky decomposition are about $M^3/3$. Furthermore, assuming the number of the observation, i.e., geopotential differences, to be N, then the FLOPs of forming a normal matrix would be about $(2\times N-1)\times(M+1)\times M/2$, or approximately $O(nm^2)$ [*Xie*, 2005]. So the total FLOPs would be the sum of the above two parts, which is $M^3/3 + (2\times N-1)\times(M+1)\times M/2$. About the memory limitation, for storing the design matrix the memory requirement is about $N\times M$. And for storing the normal matrix it is $M\times M/2$. So the total memory requirement would be $N\times M + M\times M/2$.

Figure 3.6 visualizes the complexity of the computation. Here one month GRACE data is assumed to be the input of the inversion, so N is about 518,400 (17,280 for each day because of 5 seconds sampling rate). The figures show the total FLOPs and total memory requirement with respect to the maximum degree to be solved. The top panel of Figure 3.6 shows the requirement of the FLOPs for processing one month GRACE data, where blue line represents the FLOPs for forming a normal matrix, i.e., $(2\times N-1)\times(M+1)\times M/2$, and the green line represents the FLOPs for inversing a normal matrix, i.e., $M^3/3$. The numbers in the parenthesis marks the coordinates (x,y) for selected cases. For example, if the maximum degree is 60, the requirement of FLOPs is about 7200G (mainly from forming the normal matrix, i.e., blue line). For a normal workstation with the CPU frequency of 3GHz, the computation time is about 2400 seconds, i.e., 48 minutes. The computation time can be further reduced if a parallel programming technique, such as OpenMP (Open Multi-Processing), would be considered. For instance, if the workstation has a total of 12 cores, so the computation time can be brought down to about 4 minutes if all the cores are available at the same time. For maximum degree and order larger than 60, Message Passing Interface (MPI) parallel platform from Ohio Super Computer (OSC) is utilized, which is realized by an optimized implementation of ScaLAPACK included in Intel Math Kernel Library (MKL).

The bottom panel of Figure 3.6 shows the requirement of the memeory also for processing one month GRACE data, where blue line represents the memory requirement for storing a design matrix, i.e., $N\times M$, and the green line represents the memory requirement for storing a normal matrix, i.e., $M\times M/2$. Similar to the FLOPs figure, if the maximum degree is 60, the memory requirement is about 15 GB (mainly for storing the design matrix), which is possible on a workstation. For maximum degree and order larger than 60, Design Matrix Accumulation (DMA) method is considered instead of the Normal equation Matrix Accumulation (NMA) method in order to alleviate large

memory requirement for design matrix. More information about the application of parallel computation about gravity inversion and can be found from *Xie et al.* [2004] and *Xie* [2005].

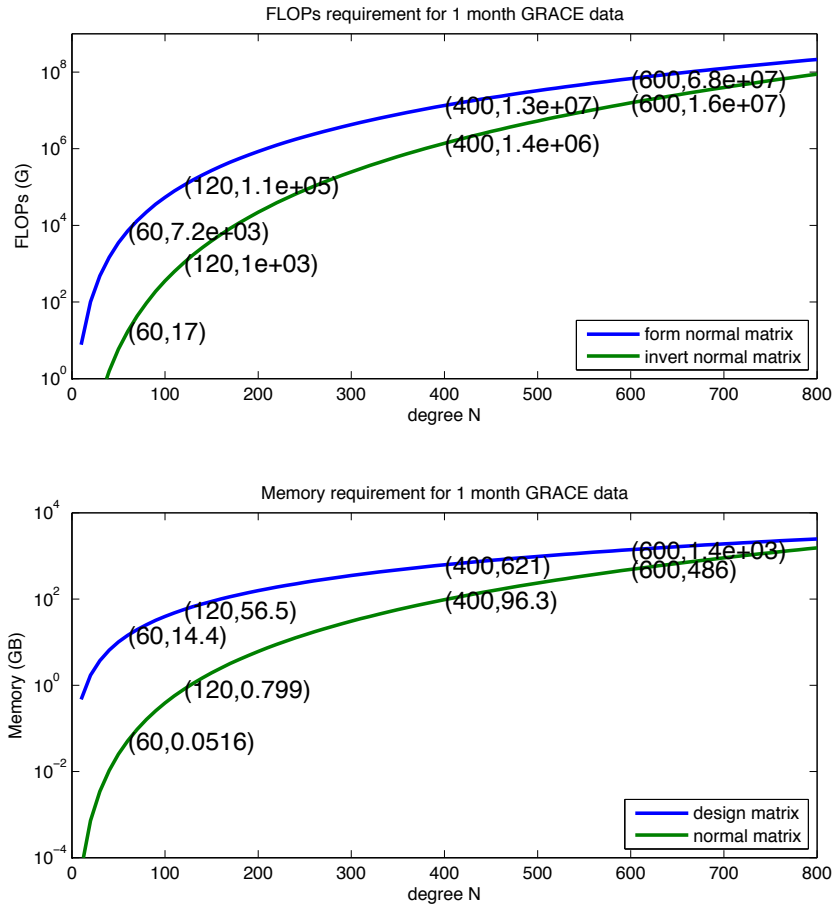


Figure 3.6 Complexity of the computation for solving a monthly global solution using 1-month geopotential difference data. Top panel: Float point operations (FLOPs) requirement. Bottom panel: Memory requirement. The numbers in the parenthesis marks the coordinates (x,y) for selected cases.

3.2.2 Recovered Monthly Gravity Solution

Using all the geopotential difference estimates from 2003 to 2013, a series of monthly global solutions is generated including most of the months during that period except a few months because of the data outage. All the global solutions are computed up to degree and order 60, which is consistent with the official L2 products. Strictly speaking, the maximum degree and order may vary depending on the number of the valid

observation during a month and more importantly the ground track distribution [*Weigelt et al.*, 2013]. As a satellite gravimetry mission, GRACE is designed to avoid repeat orbit as much as possible, i.e., with homogeneous ground track distribution. However, for a certain period, the GRACE orbit would decay into a near-repeat orbit [*Yamamoto et al.*, 2005], i.e., with overlaps and large gaps among the ground tracks, due to the orbit perturbation. It is reasonable to believe the months with homogeneous ground track distribution ('good' month) should yield better resolution of the gravity field than the months with near-repeat ground track distribution ('bad' month) [*Save et al.*, 2012]. Accordingly, if all the months are solved with an identical resolution, the gravity solution for those 'bad' months may contain larger error than the solution for the 'good' months.

Similar question is raised for the improved energy balance method. In this study, the only input for gravity inversion is geopotential differences, which are obtained directly from KBR range-rate measurements. Compared to conventional dynamic method, the GPS phase measurements are not explicitly included during the inversion. Therefore, it is important to check the KBR-only solutions for both 'good' months and 'bad' months. The example of the 'good' month here is chosen to be July 2003, and the 'bad' month is September 2004, which has a typical month of near-repeat orbits with repeat cycle of 61 revolutions in 4 days. Study by *Wagner et al.* [2006] suggests the maximum degree of monthly gravity solution of September 2004 should be around 30 instead of 60 in order to avoid the higher degree error. Therefore for some scientific applications, solutions around September 2004 were directly excluded [*Swenson et al.*, 2006].

3.2.2.1 Case with 'good' ground track coverage

Figure 3.7 shows the result for the 'good' month of July 2003. First all the estimated geopotential difference observations are shown for the month of July 2003 on the global map in Figure 3.7a. The data from descending pass are presented with an additional minus sign so they would not look opposite to the data from ascending pass over the same region of the global map. Similar to Figure 3.5, the observations can roughly indicate the surface gravity signal. As for a global map, even though they are not measured exactly at the same altitude, similar signals can still be observed when two satellites pass the same area since they were flying in similar, near-circular orbits with slight altitude decay. Therefore some regions with large gravity variation are manifested in the global map of Figure 3.7a, including not only the highlighted regions in Figure 3.5 but also some other regions like Alaska (glacier melting), Congo Basin (wet season), and Scandinavia (GIA). For comparison, the other three figures in the left panel of Figure 3.7 show the predicted values at the same geographical location from three GRACE L2 solutions (Figure 3.7c for CSR RL05, Figure 3.7e for GFZ RL05a and Figure 3.7g for JPL RL05). Again, clearly the global map of estimated data matches the global map of the predicted data from official L2 solutions very well.

The most significant discrepancy between Figure 3.7a and other three figures in the left (Figure 3.7c, Figure 3.7e and Figure 3.7g) is that Figure 3.7a apparently contains measurement noise which are inherited from each range-rate measurement, while other

three figures are only predicted from a modeled gravity field with maximum degree of 60 only. That implies probably the most significant contribution of using energy balance approach, i.e., directly connecting the geometry measurements (range-rate) to the geophysical quantities (geopotential difference). Therefore, even though the whole month estimates are shown in the same global map, Figure 3.7a still preserves the in-situ geopotential change for each epoch within a month, i.e., sub-monthly information, but Figure 3.7c, 3.7e and 3.7g can only show the predicted values from a monthly mean (static) gravity field. In next section, the possibility of extracting such sub-monthly time-variable gravity information from geopotential differences will be explored.

Using the data from Figure 3.7a, the monthly mean time-variable gravity solution up to degree and order 60 is solved and shown in Figure 3.7b. Here the Stokes coefficients are converted into geoid according to the equation

$$\Delta N = a\sum_{n=2}^{n_{\max}}\sum_{m=0}^{n}\bar{P}_{nm}(\cos\theta)\left(\Delta\bar{C}_{nm}\cos m\lambda + \Delta\bar{S}_{nm}\sin m\lambda\right), \tag{3.4}$$

where $\Delta\bar{C}_{nm}$ and $\Delta\bar{S}_{nm}$ represent the time-variable Stokes coefficient change with mean, static part removed, therefore ΔN represent the time-variable geoid change with the same mean geoid removed. The same geoid maps from other three L2 solutions for the month of July 2003 are also shown in the right panel of Figure 3.7 for comparison (Figure 3.7d for CSR RL05, Figure 3.7f for GFZ RL05a and Figure 3.7g for JPL RL05). The reason of showing geoid here instead of the common used equivalent water height (EWH) is because geoid map can highlight both signal and noise (i.e., north-to-south stripes) together without applying any post-processing. That means no post-processing techniques are applied here except the replacement of the C_{20} coefficient using the value obtained from satellite laser ranging [*Cheng et al.*, 2013] since different products show large disagreement on that coefficient.

Actually, the good agreement of the estimated and predicted data shown in left panel of Figure 3.7 has already guaranteed that my gravity solution after the inversion process, i.e., the downward continuation, should not deviate from the official L2 solutions too much. And indeed, my solution in Figure 3.7b does show similar features of signal and comparable level of noise as the other three solutions. In terms of noise level, i.e., north-to-south stripes, the geoid maps from my solution seems to be similar to the JPL solution, i.e., less stripes than the GFZ solution but more stripes than the CSR solution. The comparison of Power Spectral Density (PSD) of the coefficients from all the solutions is shown in Figure 3.8. At lower degree (below degree 15), the PSD of my solution (OSU) matches other three PSDs very well, which means my solution contains highly consistent signal of time-variable gravity. At higher degree (above degree 15), the PSD of my solution shows similar noise level as JPL's PSD, which is slightly higher than CSR's PSD but lower than GFZ's PSD.

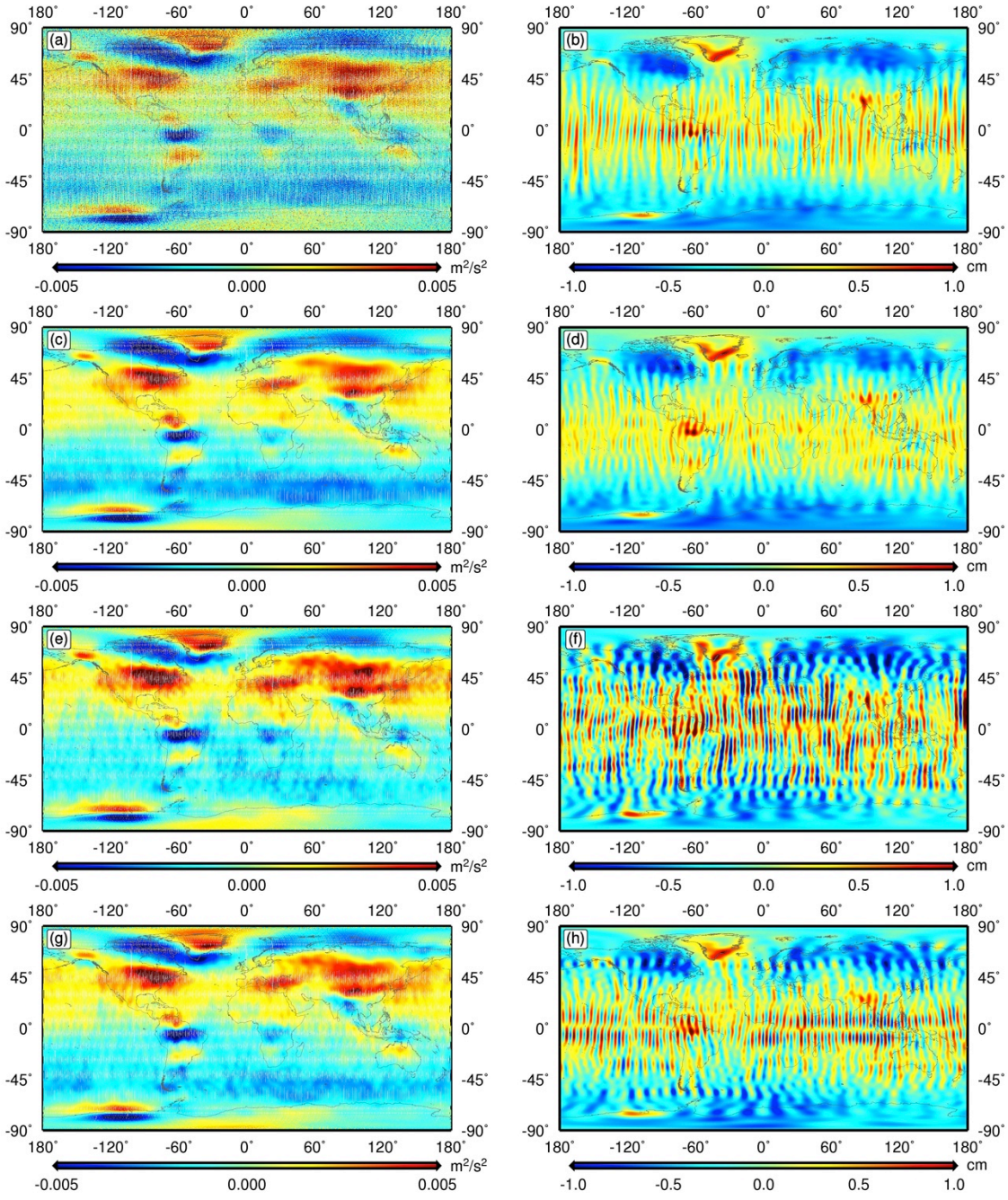


Figure 3.7 Global map of both geopotential difference estimates and recovered gravity solution (with mean field removed) for the month of July 2003. Left panel: geopotential differences (a) estimated from this study, and predicted from GRACE L2 products of (c) CSR RL05, (e) GFZ RL05a and (g) JPL RL05. Right panel: recovered geoid change from (b) this study using geopotential difference estimates and from GRACE L2 products of (d) CSR RL05, (e) GFZ RL05a, and (g) JPL RL05.

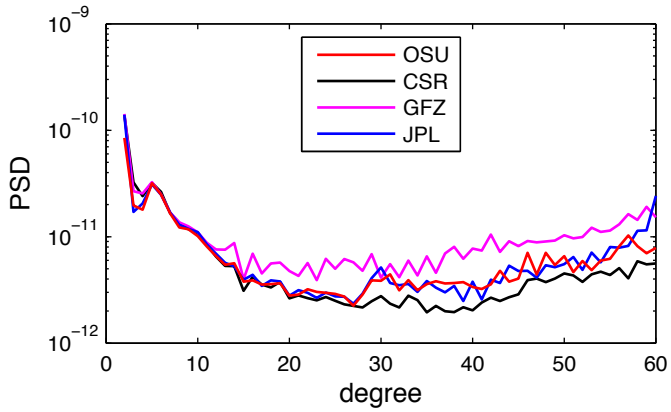


Figure 3.8 Comparison of Power Spectral Density (PSD) for the month of July 2003 from this study (OSU) and GRACE L2 products (CSR RL05, GFZ RL05a, JPL RL05).

3.2.2.2 Case with 'bad' ground track coverage

Figure 3.9 show the result for the 'bad' month of September 2004. Again, the left panel is the comparison of geopotential difference estimation (Figure 3.9a) based on range-rate and the prediction (Figure 3.9c, 3.9e, 3.9g) based on official L2 products. Compared to the left panel of Figure 3.7 for the 'good' month of July 2003, in Figure 3.9 the overlaps of the ground tracks lead to very large gaps between the geographical coverage, which is caused by the near-repeat orbit configuration during that month. GRACE should have approximate 15~16 revolutions every day because of the approximate 1.5 hours orbit period, and therefore for a normal month with 30 days non-repeat orbit, the number of the total revolution during a month should be around 450~480. However, for September 2004, because of the 4-day near-repeat period, the number of the non-repeat revolution during that month is only equivalent to 61, which is about 13% of the normal month. Therefore, more errors are expected from the recovered gravity field up to degree 60.

The solved gravity field is shown in Figure 3.9b in terms of geoid, which apparently contains heavier stripes compared to Figure 3.7b for the month of July 2003. Most of the stripes concentrate around equatorial area, which is understandable because equatorial area has worse ground track distribution than polar region. Actually for polar region, it is surprising to see that some of signals are not significantly affected by the stripes, such as Greenland. Again, the non-repeat 61 revolutions here are only equivalent to 4 days' geopotential difference observation. Therefore, it implies that shorter temporal resolution than a month might be possible, at least for certain area such as polar region.

For comparison, Figure 3.9d, 3.9f and 3.9h shows the same geoid maps from three L2 official products, CSR RL05, GFZ RL05a and JPL RL05, respectively, for the month of September 2004. Unfortunately, my result seems to contain more stripes than other three official products. It might be able to match GFZ RL05a and JPL RL05, but definitely not

comparable to CSR RL05. One of the reasons could be official L2 products include direct contribution from both GPS and KBR data by using dynamic method as mentioned before, but it is questionable how large the contribution from GPS to solution is. At least study by *Luthcke et al.* [2006] shows that monthly solution solely from range-rate measurements could achieve comparable GRACE solutions as the official L2 products, using the similar dynamic method. It is also worthy mentioning that GFZ RL05a use regularization to stabilize the solution for that month [*Dahle et al.*, 2012], which can also mitigate the stripes. It is unclear whether or not the similar strategy has been also adopted by CSR and JPL. The comparison of Power Spectral Density (PSD) of the coefficients from all the solutions is shows in Figure 3.10. Compared to Figure 3.8, similar agreement can be found at low degree. At higher degree (above degree 15), the PSD of my solution shows higher noise level than CSR and JPL's PSD, and shows similar noise level as GFZ's PDS.

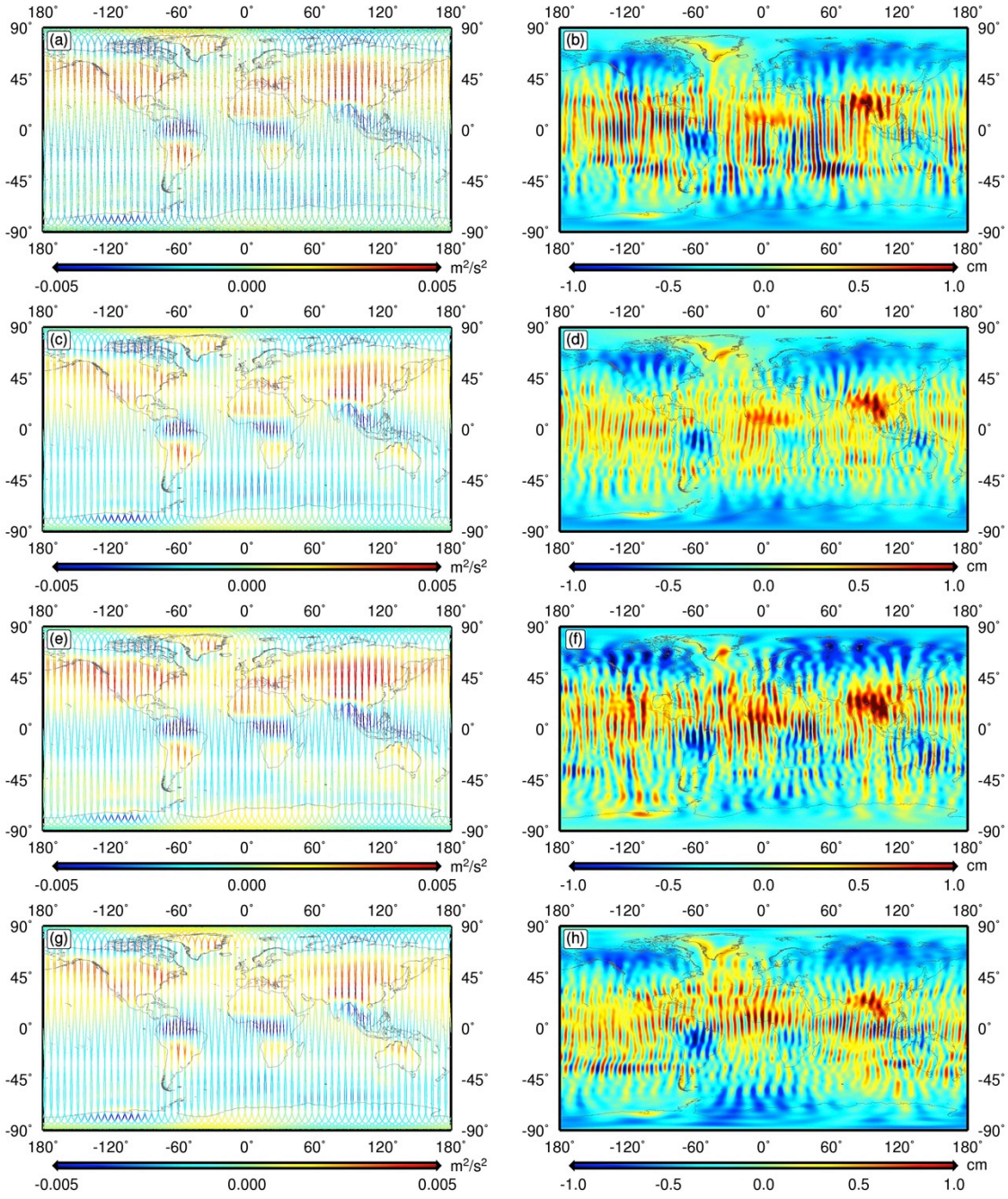


Figure 3.9 Global map of both geopotential difference estimates and recovered gravity solution for the month of September 2004. Left panel: geopotential differences (a) estimated from this study and predicted from GRACE L2 products of (c) CSR RL05, (e) GFZ RL05a and (g) JPL RL05. Right panel: recovered geoid change from (b) this study using geopotential difference estimates and from GRACE L2 products of (d) CSR RL05, (e) GFZ RL05a, and (g) JPL RL05.

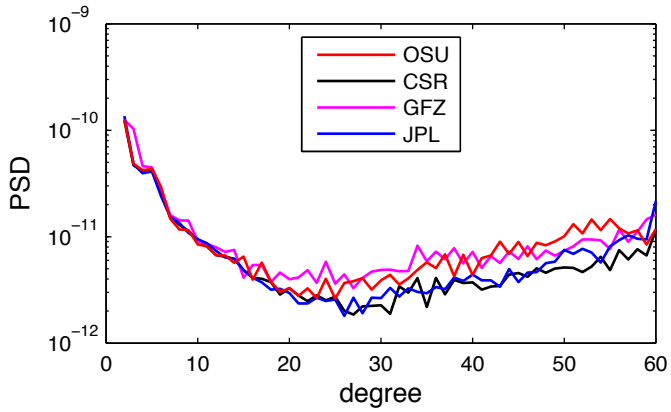


Figure 3.10 Comparison of Power Spectral Density (PSD) for the month of September 2004 from this study (OSU) and GRACE L2 products (CSR RL05, GFZ RL05a, JPL RL05).

3.2.3 Secular and Seasonal Gravity Variation From Global Solutions

As a dynamic planet, Earth's gravity is continuously varying in different time scales. The primary goal of GRACE mission is to precisely map this temporal variation, in order to better understand the dynamic process, such as mass transportation, within the Earth system. This subsection focuses on the long-term variation of the gravity, and presents the secular and seasonal variation observed by GRACE.

3.2.3.1 Secular Variation of Geoid

As said in last subsection, based on the geopotential difference estimates, a series of monthly gravity solutions are generated up to degree and order 60 from 2003 to 2013. Based on that, the secular and seasonal variations of the time-variable gravity are estimated. Figure 3.11a shows the estimated secular variation from 2003 to 2013, in terms of geoid height trend. Figure 3.11b, 3.11c, 3.11d show the geoid trend maps estimated from for CSR RL05, GFZ RL05a and JPL RL05, respectively.

Clear and consistent secular signals can be observed from Figure 3.11a-d for all the solution series, including negative trend in Greenland, Amundsen Sea Embayment, Antarctic Peninsula, and Alaska, reflecting the mass loss from the ice sheets and glacier, and positive trend in Hudson Bay, West and East Antarctic, and Scandinavia, mainly due to GIA. The agreements are mostly over high-latitude and polar region.

On the other hand, the result in this study seems to have more stripes in certain areas, especially over low-latitude or equatorial regions, such as Southeast Asia with the signal of great Sumatra-Andaman earthquake of 26 December 2004, where my trend map show heavier stripes than other three solutions. One of the reasons for the discrepancy could be the different reference model used in my processing, especially the high-frequency

models that could regionally impact the aliasing effect if the model is less accurate over certain area. Another reason may be caused by the 'bad' months as I discussed in last subsection. The monthly solutions from 'bad' months should contain more aliasing errors, i.e., stripes, than the 'good' months. The geoid estimated from the 'bad' months may significantly contaminate the trend estimation. Energy balance method may suffer this problem more than the dynamic method because of the absent of the direct contribution from GPS. The strategy of applying regularization on those 'bad' months (at least by GFZ) might also be a reason.

Figure 3.12 shows the comparison of Power Spectral Density (PSD) of the trend field from all the solutions. All the PSD agree with each other before around degree 25. After that, PSD of my solution deviates from CSR and JPL's PSD, but is close to GFZ's PSD. JPL's PSD deviates from CSR's after around degree 40. PSD of my solution derivate from GFS's and drops around degree 56, but normally signals beyond about degree 55 are regarded as noise for GRACE. The reason of the similarity between my PSD and GFZ as well as that between CSR and JPL might be also caused by the different reference models used during the data processing, since the models I adopt are identical to the models used by GFZ [*Dahle et al.*, 2012], which is different from the models shared by JPL [*Watkins et al.*, 2012] and CSR [*Bettadpur*, 2012], especially for ocean tides models as well as the interpolation method of minor ocean tides.

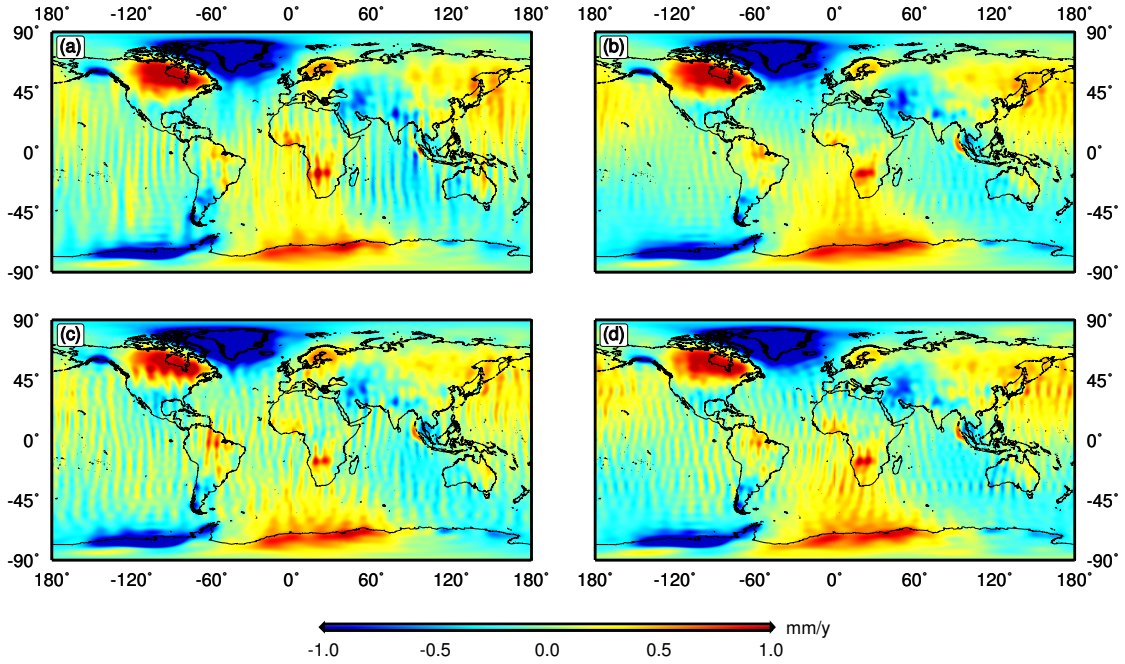


Figure 3.11 Geoid trend map (2003~2013) from this study (OSU) and official products. (a) OSU, (b) CSR RL05, (c) GFZ RL05a and (d) JPL RL05.

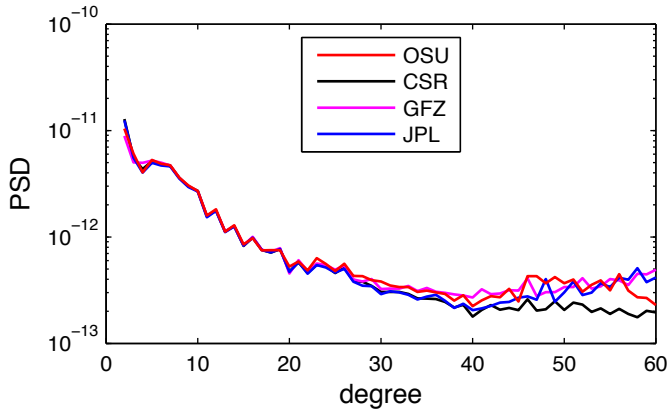


Figure 3.12 Comparison of Power Spectral Density (PSD) for estimated trend from 2003 to 2013 from this study (OSU) and GRACE L2 products (CSR RL05, GFZ RL05a, JPL RL05).

3.2.3.2 Secular Variation of Equivalent Water Height

A large quantity of the application using GRACE monthly solutions is to estimate the surface mass transportation for hydrology and glaciology study. Therefore, it is necessary to build a connection between the surface mass transportation and gravity or geoid change. *Wahr et al.* [1998] first solved the problem based on a thin layer assumption, i.e., all the GRACE observed gravity change is caused by the surface density change of a thin layer as well as the corresponding loading effect on solid Earth underneath. The thin layer is on the reference sphere defined by the mean radius of the Earth, i.e., 6378136.3 meters, and the thickness of the layer is negligible compared to the radius of the Earth, but has to be large enough to include all the mass change from atmosphere, cryosphere, ocean, terrestrial and ground water which are on the order of 10–15 km. The derived surface density change can be further scaled by the water density and converted into equivalent water height (EWH) change as

$$\Delta h = \frac{a\sigma_E}{3\sigma_w}\sum_{n=0}^{\infty}\frac{(2n+1)}{(k_n+1)}\sum_{m=0}^{n}\bar{P}_{nm}(\cos\theta)\left(\Delta\bar{C}_{nm}\cos m\lambda + \Delta\bar{S}_{nm}\sin m\lambda\right), \tag{3.5}$$

where a is the Earth's mean radius, σ_E is the average density of Earth, σ_w is the density of water, k_n are Love numbers of degree n. In Figure 3.13, the geoid trend map is converted into the EWH trend map based on the equation (3.5). Unlike the geoid map shown in Figure 3.11, the EWH map in Figure 3.13 seems to contain heavier stripes, which is caused by the amplification of high-frequency noise due to the additional ($2n$+1) factor in equation (3.5). The trend map from CSR shows fewest stripes than others. Trend maps from other three solutions shows similar heavy stripes but with difference characteristic. Trend map from this study show more stripe near equatorial region but less over polar region than either GFZ or JPL solution. Especially over Greenland and Antarctica, the trend map from my solution manifests similar noise level as the CSR solution.

Therefore, post-processing techniques are required in order to mitigate those error and noise. The south-north stripes are caused by the correlation between the orbit-sampling period and certain gravity coefficients, mainly around the resonant orders of 15~16, 30~32 and 45~48. *Swenson et al.* [2006b] provided an ad hoc method to reduce the error. Thereafter, similar (e.g., *Chambers*, [2006]; *Chen et al.* [2007]; *Duan et al.*, [2009]) as well as alternative (e.g., *Davis et al.*, [2008]; *Klees et al.*, [2008]; *Kusche*, [2007]; *Schrama et al.*, [2007]; *Wouters et al.*, [2007]) approaches have been devised to reduce the correlation error. Here, the method from *Duan et al.* [2009] is adopted as a 'de-correlation' technique. Besides de-correlation technique, Most of the approaches also need a further spatial domain smoothing (e.g., *Guo et al.*, [2010]; *Han et al.*, [2005]; *Jekeli*, [1981]; *Wahr et al.*, [1998]). The method from *Wahr et al.* [1998] is adopted as smoothing technique, which is same as the isotropic Gaussian filter from *Guo et al.* [2010].

Figure 3.14, 3.15 and 3.16 show the same EWH trend maps as Figure 3.13, but with difference post-processing technique applied. Figure 3.14 shows the EWH trend maps with only Gaussian smoothing applied, where the average radius is 200 km. Apparently, Gaussian smoothing alone is not enough to reduce all the stripes on the trend map, at least for the solution from this study and also GFZ and JPL. CSR shows an overall best result, so it might be able to use only Gaussian smoothing without de-correlation for CSR solutions. Figure 3.15 shows the EWH trend maps with only de-correlation applied. After de-correlation, my result shows similar noise level as other three official products, which means correlated error mainly dominates my result (as well as GFZ and JPL) instead of white noise. Figure 3.16 shows the EWH trend maps with both 200 km Gaussian smoothing and de-correlation applied. As a result, the combination of de-correlation and 200 km Gaussian smoothing can effectively reduce both the correlated error and white noise, and lead to almost identical trend map from my solution and other three official products.

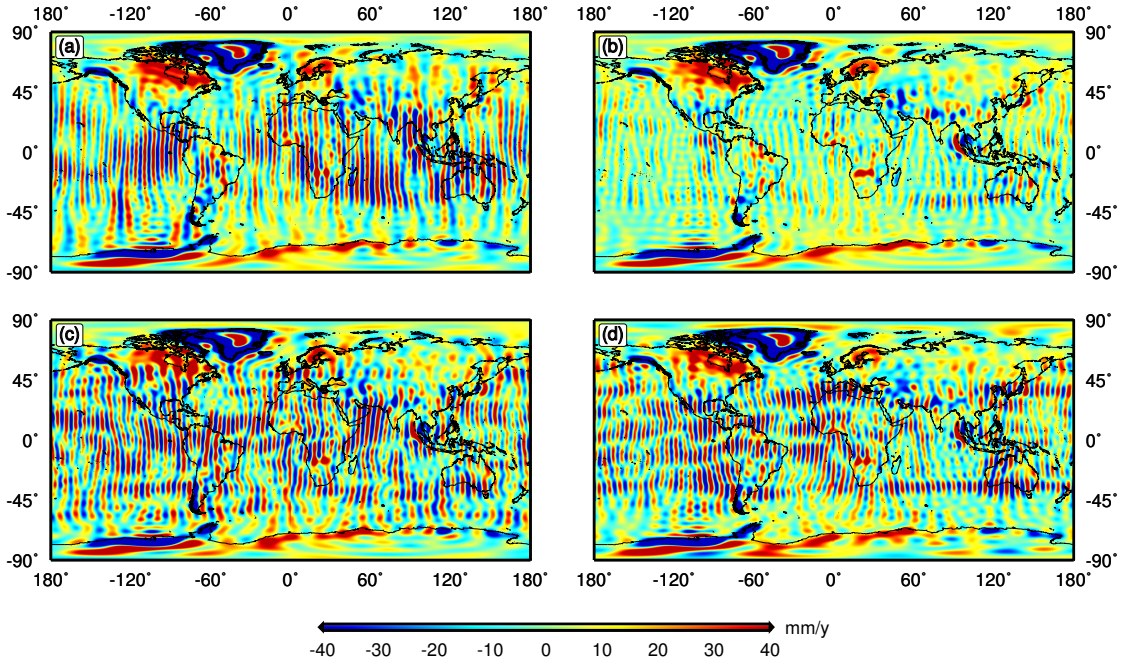


Figure 3.13 Equivalent Water Height (EWH) trend map (2003~2013) from this study (OSU) and official products. (a) OSU, (b) CSR RL05, (c) GFZ RL05a and (d) JPL RL05.

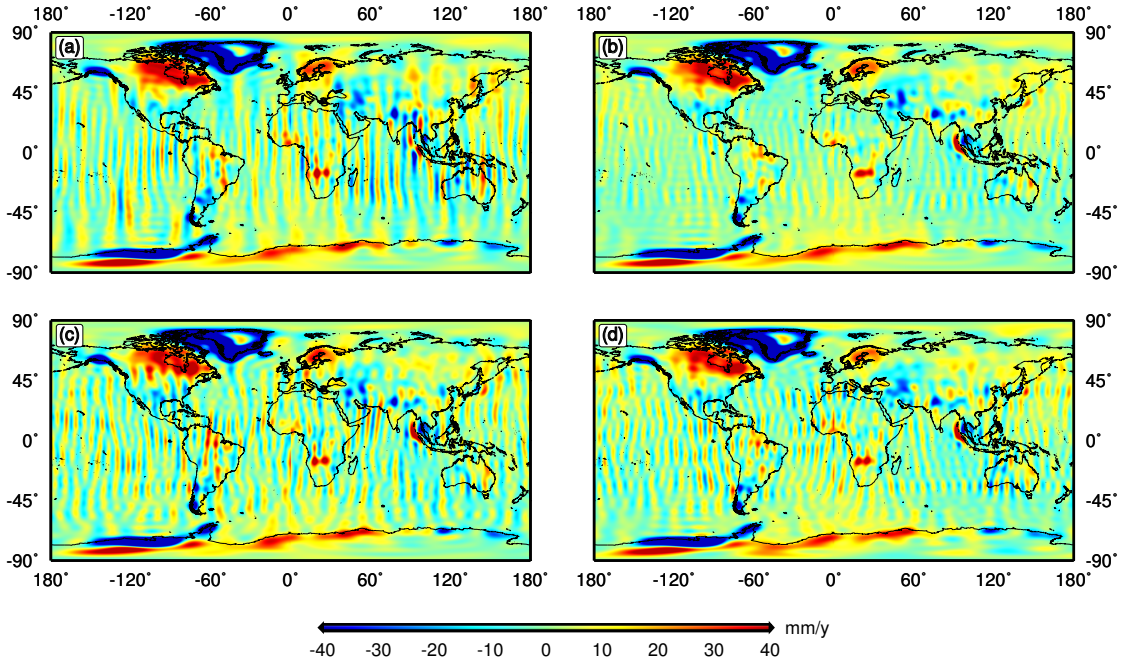


Figure 3.14 Same as Figure 3.13 except 200km Gaussian smoothing applied.

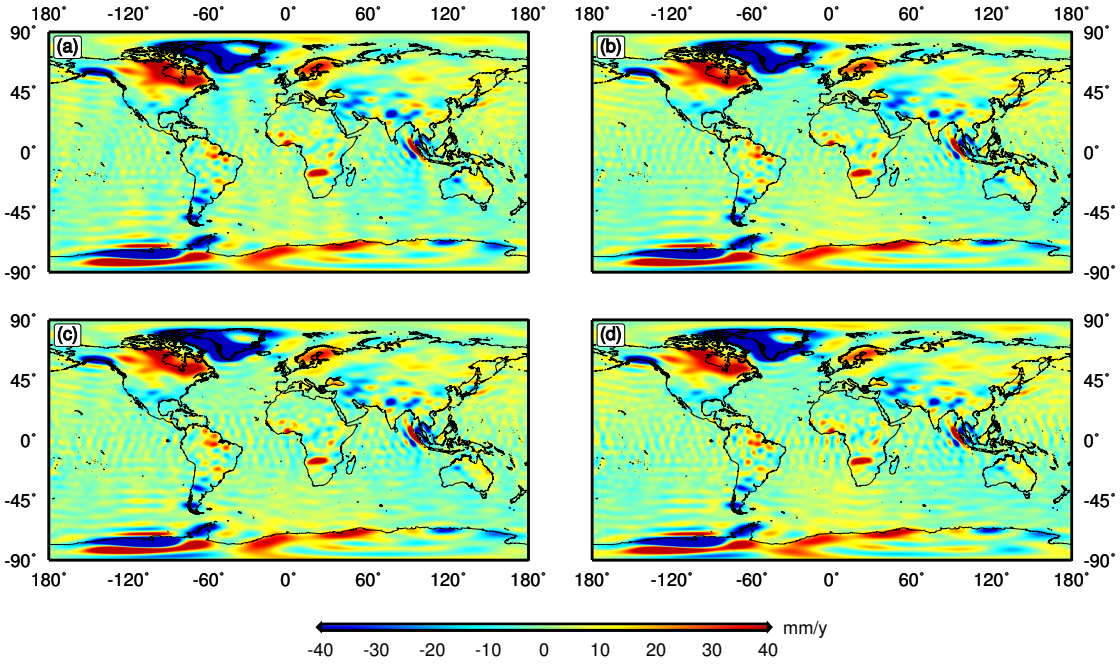


Figure 3.15 Same as Figure 3.13 except de-correlation applied

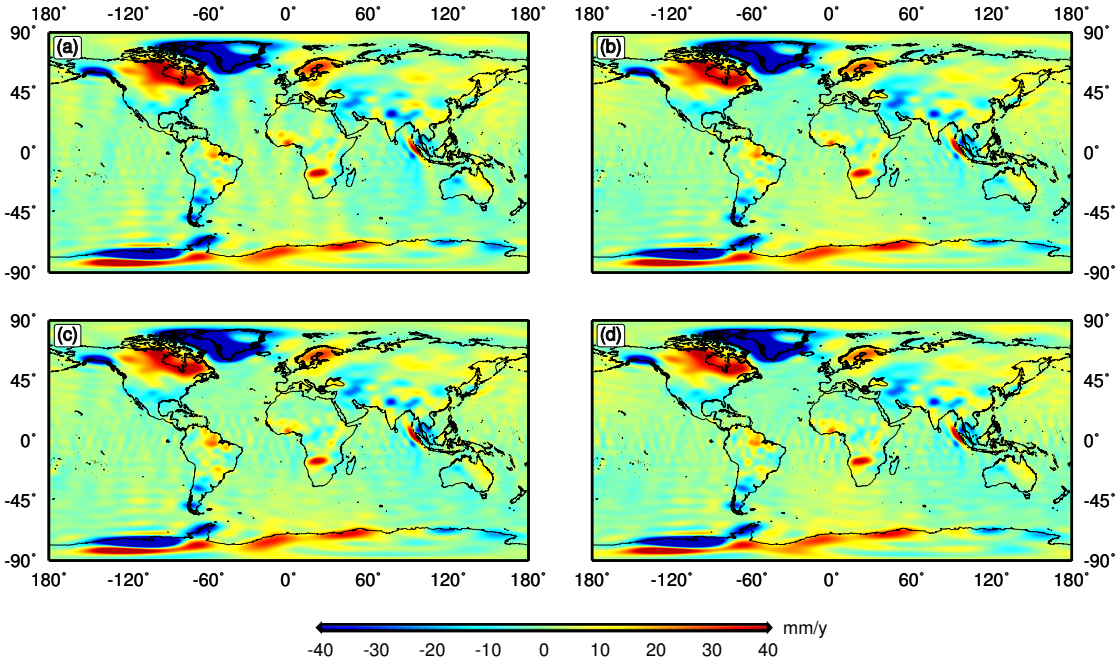


Figure 3.16 Same as Figure 3.13 except de-correlation and 200km Gaussian smoothing applied.

3.2.3.3 Seasonal Variation over Selected Basins

Trend map is an efficient way to evaluate the secular change. But in order to show the seasonal variation in addition to secular variation, the monthly solutions are still needed to show the seasonal change of the EWH. Here several regions are chosen to generate the time series in terms of EWH using all the monthly solutions from 2003 to 2013. Similar to the trend map in Figure 3.16, both 200 km Gaussian smoothing and de-correlation are applied to each monthly solution. 33 world major river basins are chosen as the study regions to compute the mean EWH variation. The boundaries of those river basins are given in *Llovel et al.* [2010] as shown in Figure 3.17, with the corresponding basin ID from 1 to 33. Besides that, two major ice sheets, Antarctica and Greenland, are also included, with basin ID of 34 and 35, respectively.

The times series of mean EWH from 33 river basins as well as 2 ice sheets are illustrated from Figure 3.18 to Figure 3.20. For river basins, these time series reveal the seasonal cycle for most of the basins, and usually the larger the basin is, the smoother the time series looks like. The GRACE measured EWH represents the total water storage (TWS), which include the surface water, soil moisture and ground water. For some basins like Amazon (basin ID 01), surface water redistribution and soil moisture change dominate the observed variation, which is mainly caused by precipitation, evapotranspiration and runoff, indicating GRACE could serve as a powerful tool to help us understand the dynamics of the surface hydrology process [*Han et al.*, 2009]. For some region like Indus (basin ID 13), the seasonal cycle is driven by both surface water and ground water change. The declined trend signal indicates the ground water depletion over the Indus River plain aquifer [*Rodell et al.*, 2009].

For the time series of the two ice sheets, a clear decline signal indicates the mass loss from both Antarctica and Greenland. The seasonal signal is due to the ice accumulation and ablation. But unlike river basin, the time series here are largely contaminated by the leakage error, which is induced by the Gaussian smoothing, and should be corrected by an additional reduction (e.g., *Baur et al.*, 2009; *Chen et al.*, 2006; *Guo et al.*, 2010; *Swenson and Wahr*, 2002). Because this correction is omitted, the amplitudes here are actually underestimated, especially for the two ice sheets.

For the inter-comparison of different GRACE solution, the time series from our study shows an overall good agreement with other three time series for both annual amplitude and phase. Slight disagreement can be identified after 2013, such as Amar (basin ID 01), Eyre (basin ID 11), Nelson (basin ID 19) and St Lawrence (basin ID 27), which will be left for further study.

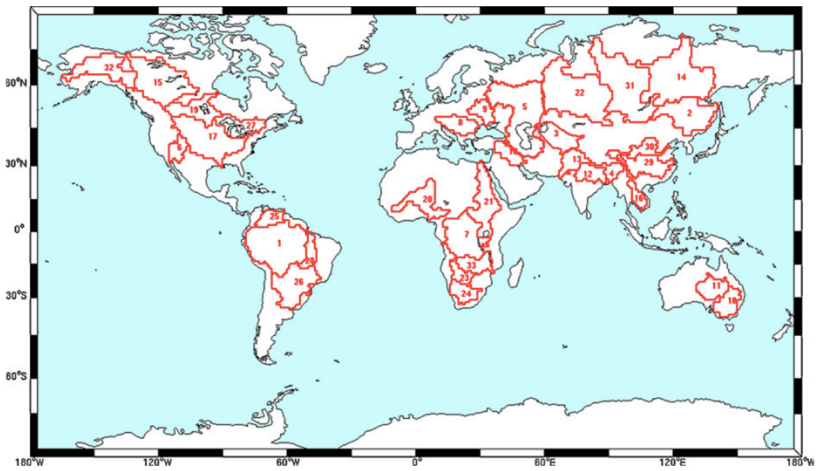


Figure 3.17 The boundaries of 33 major river basins and the corresponding basin ID given by *Llovel et al.* [2010]

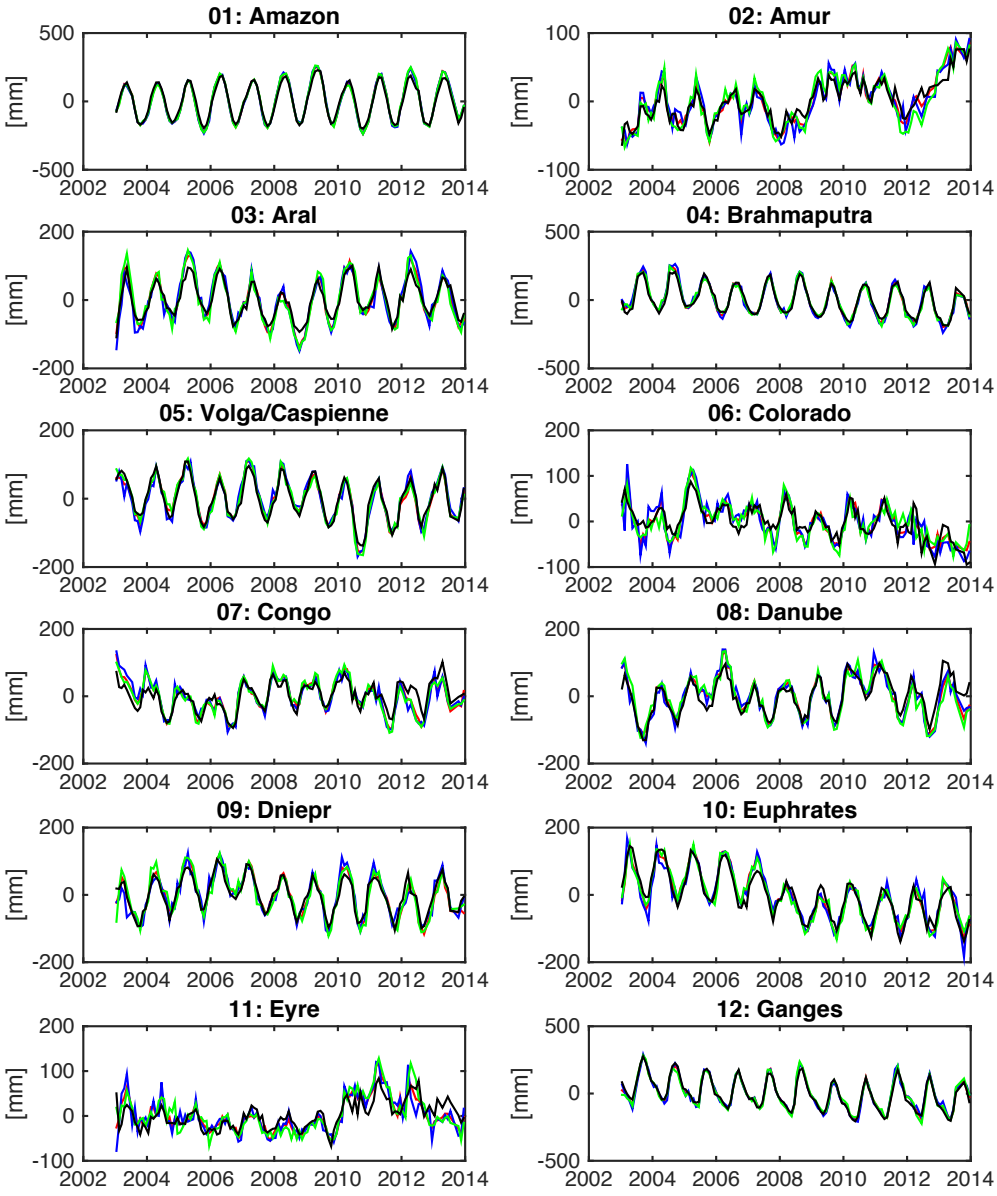


Figure 3.18 The mean EWH time series for river basins (ID 01~12), computed using GRACE monthly solutions from this study and other three official products.

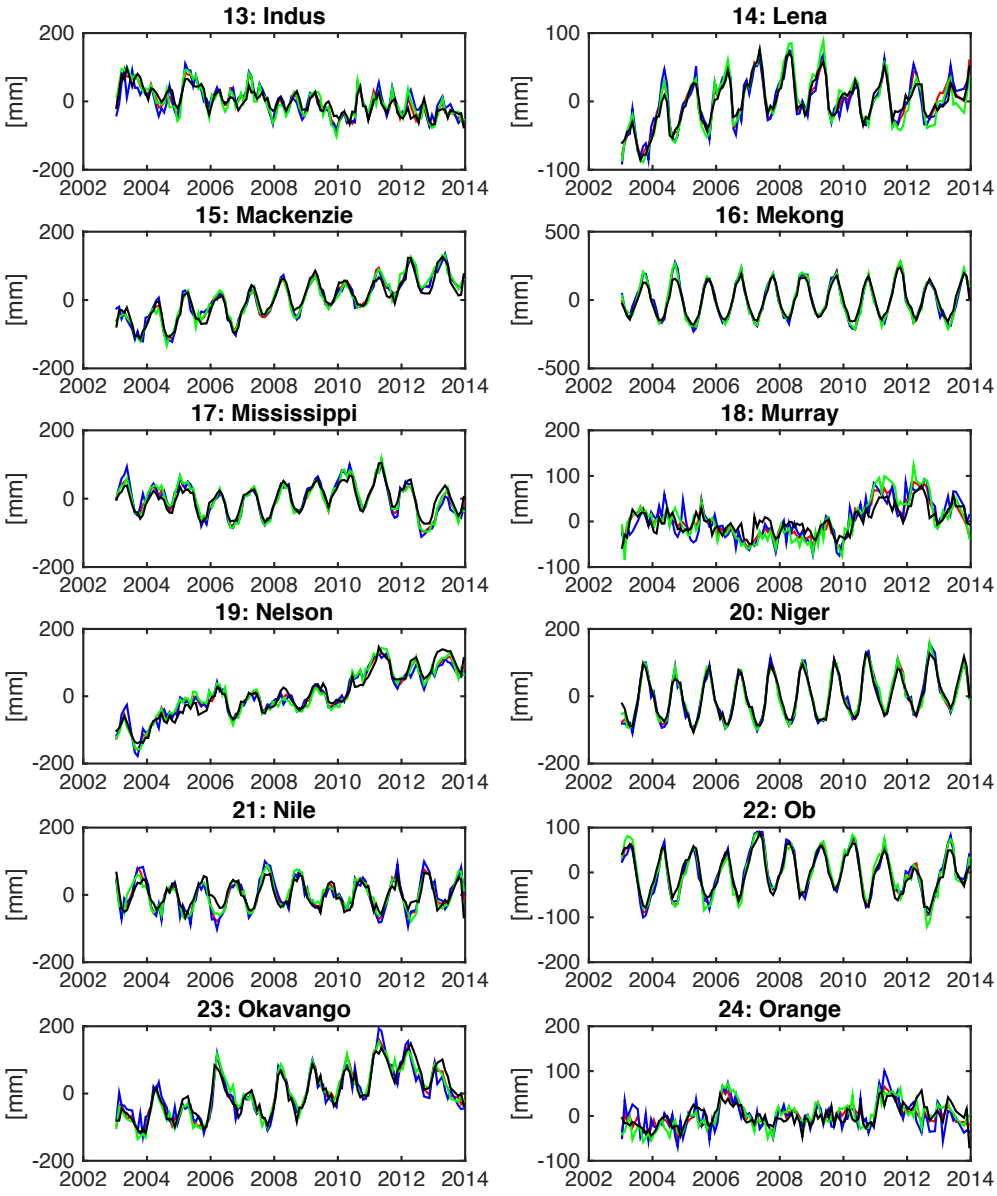


Figure 3.19 The mean EWH time series for river basins (ID 13~24), computed using GRACE monthly solutions from this study and other three official products.

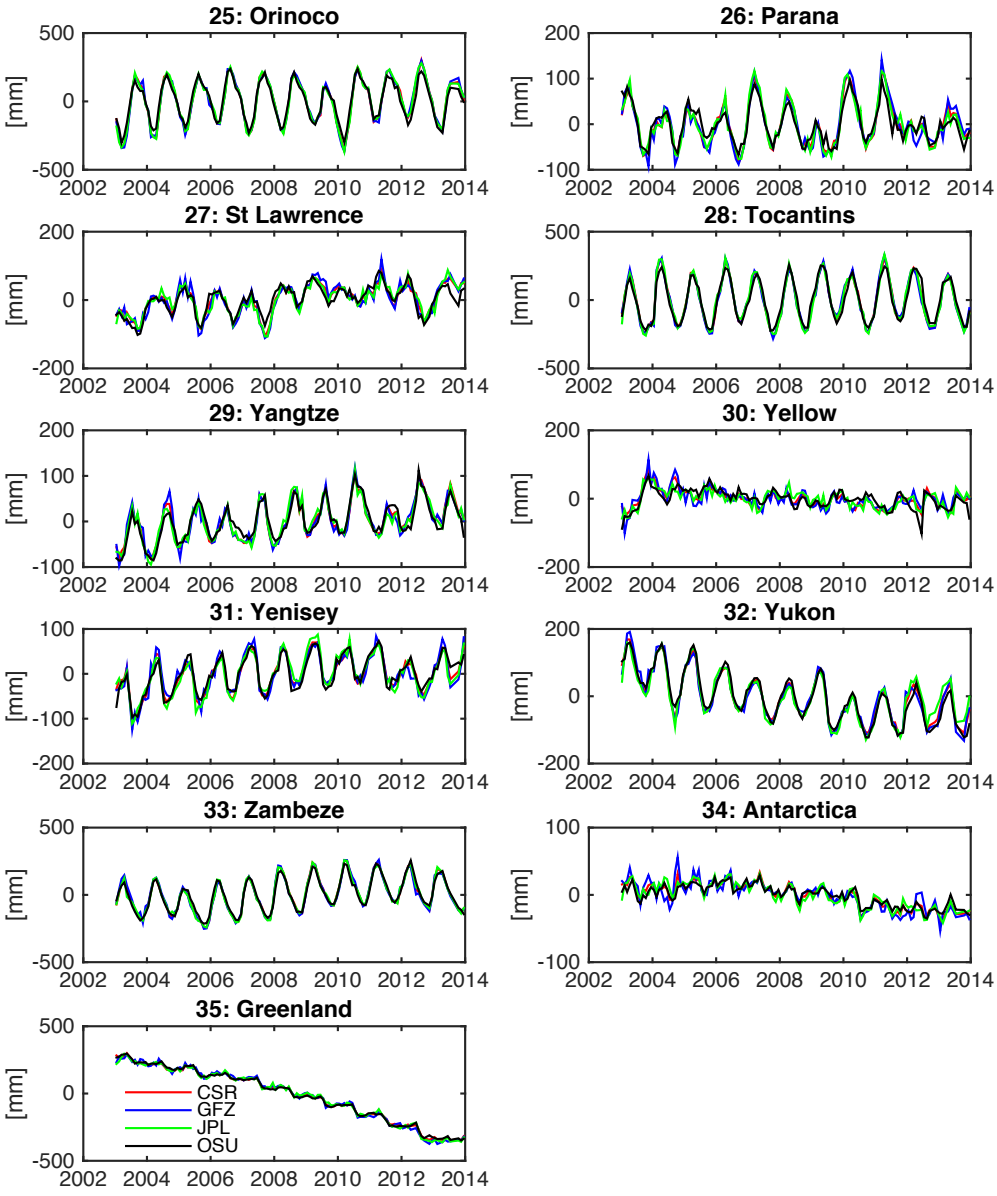


Figure 3.20 The mean EWH time series for river basins (ID 25~33) and two ice-sheet over Antarctica and Greenland, computed using GRACE monthly solutions from this study and other three official products.

3.3 Applications of GRACE Solutions with Enhanced Temporal Resolution

On a monthly basis, GRACE has been routinely measuring the Earth's time-variable gravity field from space at approximate 400~500 km altitude since late 2002 with only a few gaps. There exists a tremendous amount of publications about hydrology studies using GRACE, from the monitoring of flood [*Chen et al.*, 2010] and drought [*Chen et al.*, 2009], to the assessment of hydrological models [*de Paiva et al.*, 2013] and separate components (such as groundwater [*Döll et al.*, 2014], precipitation [*Matsuo and Heki*, 2012], evapotranspiration [*Rodell et al.*, 2004] and runoff [*Han et al.*, 2009]). Most of these studies start with the estimation of Total Water Storage (TWS) [*Wahr et al.*, 1998] using the L2 monthly solutions.

However, it is rational to presume a relative finer or enhanced temporal resolution from GRACE. First, as it can be seen from Figure 3.7 and 3.9, even though the ground track distribution for the 'bad' month (Figure 3.9) is much worse than that for the 'good' month (Figure 3.9), the resulting gravity solution in Figure 3.9 can still yield valid time-variable gravity signal. Secondly, global gravity recovery usually has a homogeneous resolution by truncating the solution to a maximum degree (and order), such as degree 60 for RL05, corresponding to the spatial resolution of 333 km. But it is reasonable to expect the region with more data coverage should have a better resolution than the region with less data coverage. Considering the near-polar orbit configuration, it is obviously that GRACE data should have more coverage over high-latitude region than middle- and low-latitude region. Therefore, it might be also possible to obtain enhanced spatial and/or temporal resolutions over certain region, especially near polar region, using geopotential difference observation.

Indeed, several attempts have also been made to improve the temporal resolution of GRACE. For example, GFZ routinely generates weekly solutions, but only up to degree and order 30. *Kurtenbach et al.* [2012] employed short-arc method [*Mayer-Gürr et al.*, 2007] under the principle of Kalman smoothing to conduct the daily snapshot solution, however the stochastic behavior of the gravity field has to be considered as a priori information. *Kang et al.* [2008] used traditional dynamic method to generate the so-called 'quick-look' solution with a moving-window strategy (with a window step of one day and window width of 15 days), but those solutions are stabilized using regularization. An incomplete list of GRACE solutions with various temporal resolutions from different research groups can be found at http://icgem.gfz-potsdam.de/ICGEM/TimeSeries.html.

The improved energy balance approach can also be easily implemented to generate solutions with enhanced temporal resolution since the inversion using geopotential differences is very straightforward. In this section, the possibility of solving GRACE solutions with enhanced temporal resolution is investigated, and the results are applied to three different cases. The first one focuses on the Greenland ice sheet. The second one is for the terrestrial and ground water storage study over North China Plain aquifer. The

third one is for terrestrial water storage study for the case of the 2008 Southeast China snow and ice storm.

3.3.1 Sub-monthly Geoid Change over Greenland

Unlike other sub-monthly solution, I attempt to solve unconstraint solutions without the compromise of spatial resolution (i.e., degree and order 60), nor the need to use regularization or a priori information. After several empirical experiments, I found for most of the case, using 10~11 days' geopotential differences would be able to solve a solution up to degree 60 and meantime can avoid significantly increasing the correlated error, i.e., stripes, especially over polar region.

Here Greenland ice sheet is chosen as the test area. First half year's geopotential difference data from July 2013 to December 2013 are collected, and for each month the data are divided into three separate data subsets for every ten or eleven days. Then using the data from each 10 or 11-day subset, gravity solution is solved up to degree and order 60 without any constraint. In Figure 3.21, both the geopotential difference data and the corresponding solutions in terms of geoid change are presented, from top to bottom for the month from July 2003 to December 2003, respectively. Here no post-processing techniques are applied to any of the solutions. The left three columns of Figure 3.21 show the 10 or 11-day geopotential difference subsets in terms of data coverage map over Greenland, where the color represents the value with mean field removed; the middle three columns show the 10 or 11-day solutions up to degree 60 in terms of geoid undulation with mean field removed; and for comparison the right three columns show the corresponding monthly solutions also up to degree 60 in terms of geoid with mean field removed, from this study, CSR RL05 and GFZ RL05a, respectively.

It is interesting to note that although the temporal resolution is shortened to about 10 or 11 days, the data coverage is still fairly dense over Greenland for most of the subsets. And the resulting 10-day solution, in the middle three columns of Figure 6, shows explicit geoid variation, surprisingly without introducing more stripes. The existing of some 'gaps' in the data coverage, such as the third subsets for August and September, due to missing days for each month (day of year 240 and 270), do not increase the stripes over the recovered solution either. It is thus remarkable to realize that, for the region like Greenland, 10 or 11-day geopotential difference data are sufficient to recover the time-variable gravity, with a resolution up to degree 60 but without significantly increasing the error. My 10 or 11-day solutions show almost the same level of the stripes as the monthly solutions from both this study (column 7 in Figure 3.21) and CSR (column 8 in Figure 3.21). For comparison, the monthly solutions from GFZ (column 9 in Figure 3.21) even show larger stripes than my 10 or 11-day solutions during the same period.

From my 10-day solutions, the geoid fluctuation within each month can be clearly observed through the three sub-monthly geoid maps. For example, the three geoid maps of July show apparently deceasing signal from the beginning to the end of the month, and

on October geoid maps show increasing signal within that month. Also the average of the three sub-monthly geoid maps seems analogous to the monthly mean geoid map.

Therefore, I conclude that my method and the resulting geopotential difference data can substantially improve the temporal solution of the regional gravity recovery, which is conducted over Greenland ice sheets and yields an enhanced temporal resolution by 3 times. Furthermore, these 10 or 11-day solutions over Greenland might also be able to reveal more detailed sub-monthly temporal mass variations for the Greenland ice-sheets. More details evaluation using surface mass balance model will be left for future studies.

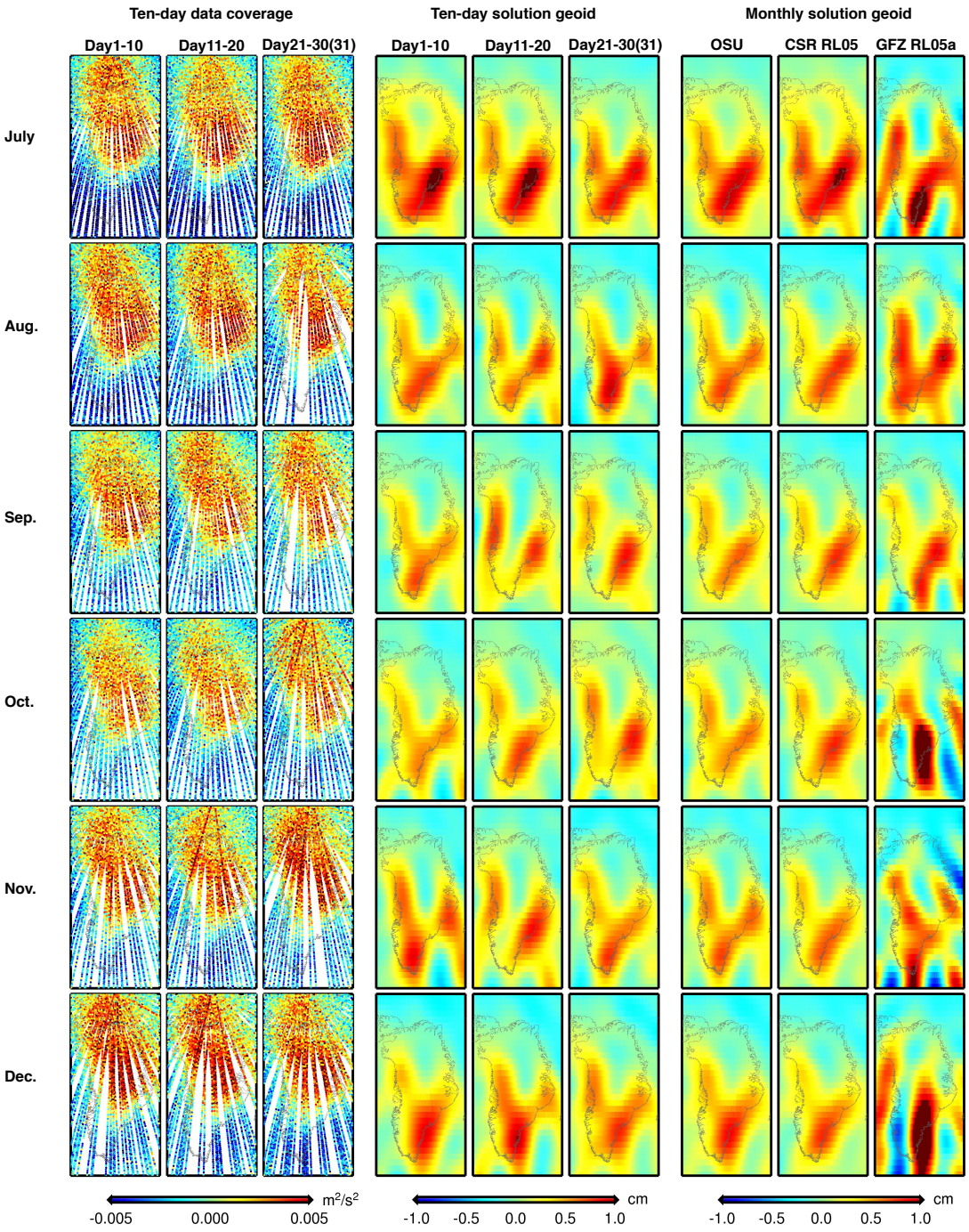


Figure 3.21 Regional enhanced solutions from July 2003 to December 2003 over Greenland ice sheet. Left three columns: 10-day geopotential difference subset in terms of data coverage map, where the color represents the value with mean field removed. Middle three column: 10-day solutions up to degree 60 in terms of geoid undulation with mean field removed. Right three columns: monthly solutions up to degree 60 in terms of geoid with mean field removed, from this study, CSR RL05 and GFZ RL05a, respectively. Neither regularization nor post-processing is applied.

With the confidence of the valid signal in my 10 or 11-day solutions, I apply the same method to solve gravity field but with only 1-day interval. That means instead of solving 3 independent sub-monthly solutions within a month, I solve time-variable gravity field for each day using the neighboring 11-day geopotential difference data, i.e., 5 days before and 5 days after the corresponding day. Those 11-day solutions with 1-day step are not independent with each other any more, and represent 11-day average solutions.

In the next two subsections, I use these 11-day solutions with 1-day step to study the sub-monthly change of the terrestrial water storage, for two cases of snow and ice storm and ground water storage change. Since the solutions need to be converted into EWH, additional post-processing techniques are inevitable. Here the routine post-processing techniques are applied for all the solution, including both de-correlation and 200 km Gaussian smoothing.

3.3.2 Terrestrial and Ground Water Storage Variation over North China Plain

This subsection presents the study of Terrestrial Water Storage (TWS) as well as Ground Water Storage (GWS) variation over North China Plain (NCP) from the GRACE. First, the monthly global solutions based on improved energy balance method are used to generate the long-term trend estimation of both regional TWS and GWS. Similar to previous studies based on official monthly solutions, my result reveals the same negative trend over this area, which may be related to the groundwater depletion in NCP. Next, six months' geopotential difference data are used for the sub-monthly analysis. Again, the strategy of 11-day solutions with 1-day step is applied here, and the solutions are solved up degree and order 60 using 11 equally weighted days of geopotential differences, which are estimated directly from Level 1B data by the improved energy balance method. The resulting sub-monthly solutions, after being routine post-processed, are compared with both hydrology models and in-situ observation.

3.3.2.1 Regional Trend Comparison

Before presenting the sub-monthly solutions, it is necessary to confirm the consistency between my monthly product and other three official products. The monthly solutions from 2003 to 2010 are used to compute the trend in terms of TWS and show the spatial pattern over China area in Figure 3.22, as well as the trend maps estimated from other three official products. Here the routine post-processing techniques are applied for all the monthly solution, including both 200 km Gaussian smoothing and de-correlation. Figure 3.22a exhibits the TWS trend estimated from this study using the monthly solutions based on geopotential difference data, referred to as OSU solution here. Some significant signals are revealed by GRACE over this area, which has been thoroughly discussed by the previous studies. The largest negative signal is over North Indian and Himalayas area, which are believed due to the groundwater depletion [*Rodell et al.*, 2009] and glacier melting [*Yi and Sun*, 2014], respectively. Another negative signal is near Southwest China, which corresponds the Tienshan glacier melting [*Yi and Sun*, 2014]. In Northeast

China, there is also a clear negative trend, reflecting the terrestrial water storage change. Figure 3.22b~d shows the trend estimated using the monthly solutions from CSR L2 RL05 product, GFZ L2 RL05a product, and JPL L2 RL05 product, respectively. By comparing my result with other three official results, it can be seen that the trend estimated using my product is highly consistent with other three products. The signals over the same location with the similar amplitude are manifested.

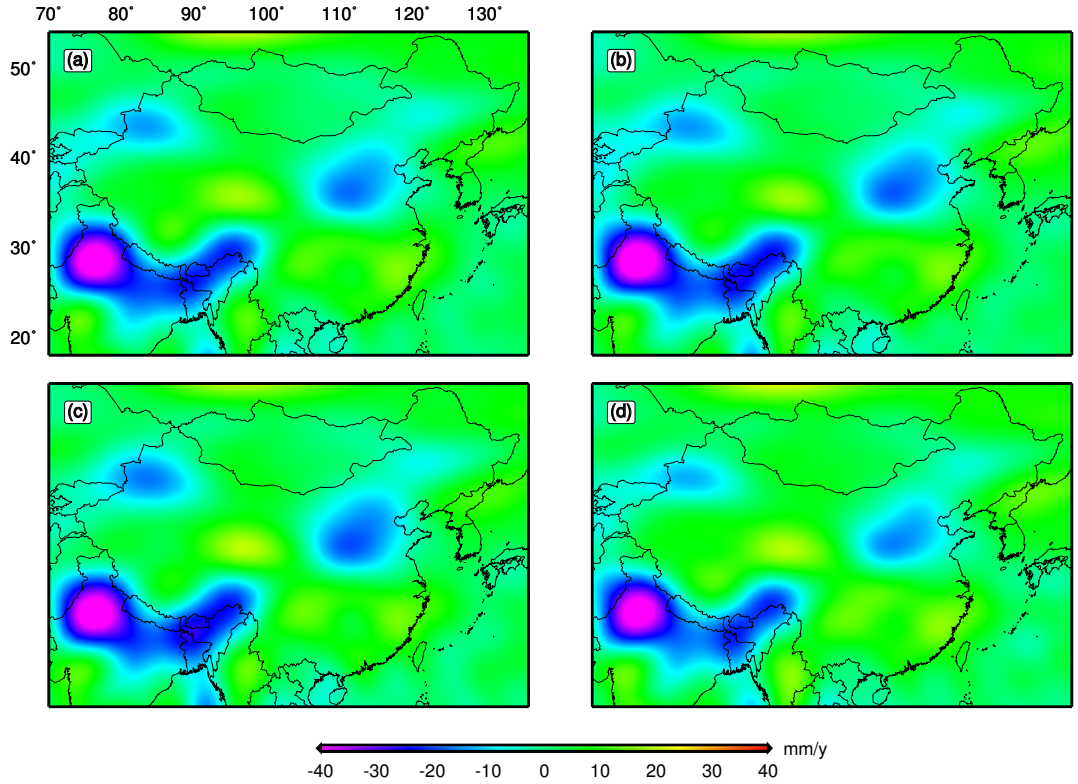


Figure 3.22 Terrestrial water storage (TWS) trend map from 2003 to 2010 over China and surrounding regions. De-correlation and 200 km Gaussian smoothing applied. Negative trend over Tienshan glacier, Indian and Himalayas area, and North China Plain. (a) from this study, (b) from CSR L2 RL05, (c) from GFZ L2 RL05, (d) from JPL L2 RL05.

Figure 3.22 (a) and (b) are almost identical to the one from a previous independent study [*Feng et al.*, 2013], which concludes the negative trend in NCP may be caused by the groundwater depletion. *Feng et al.* [2013] adopted GRACE RL05 to compute TWS, and then removed the soil moisture (SM) as well as the snow depth based on hydrological models to obtain GWS, i.e., GWS = TWS – SM. For the SM part, they used the average values of four hydrological models, three GLDAS models (NOAH, VIC, and MOSAIC) and the Climate Prediction Center (CPC) model (from NOAA). Here I apply the similar method (same smoothing but difference de-striping method) to the same data (CSR RL05) and obtain the similar results in Figure 3.23. Figure 3.23 left shows the TWS trend

over NCP, and the right shows the GWS trend based on the SM from the average of the four hydrological models. Compared to Figure 5 (a) from *Feng et al.* [2013], Figure 3.23 left shows similar spatial pattern. Here the white mask shows the administrative boundaries of four nearby provinces (Hebei and Shanxi) and municipalities (Beijing and Tianjing), and the black mask shows the hydrogeological boundary of NCP aquifer. It is noticeable that the center of the negative trend of the GRACE GWS (the dark blue dome) is not exactly located over the NCP area, but instead it is shifted to the west edge of the aquifer, close to the Taihang Mountain. *Huang et al.* [2015] further argued that this dome shift might be caused by the faster decline in the west (Piedmont Plain) than in the East Central Plain.

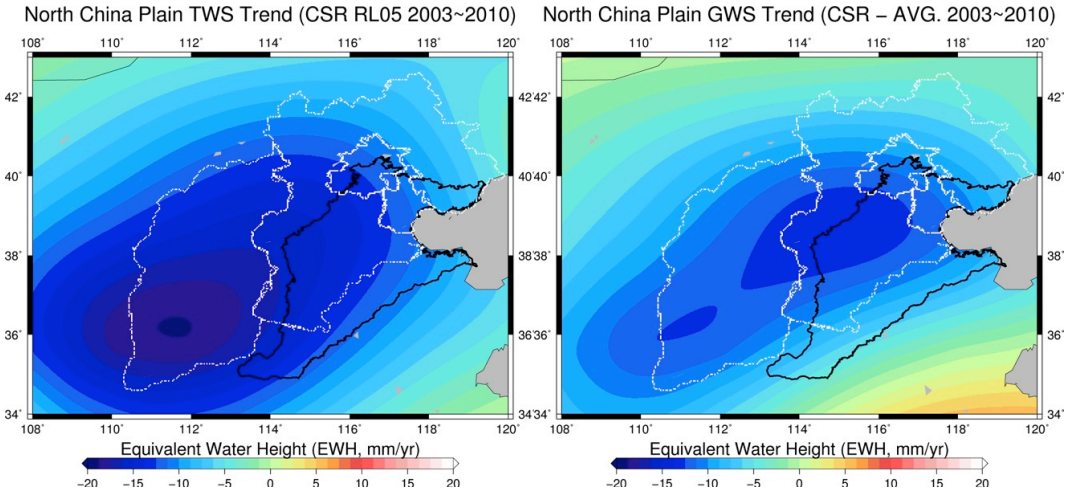


Figure 3.23 Terrestrial Water Storage (TWS) and Ground Water Storage (GWS) trend map from 2003 to 2010 over North China Plain. De-correlation and 200 km Gaussian smoothing applied. Left: TWS directly from GRACE. Right: GWS after removing average SM of four hydrological models (NOAH, VIC, MOSAIC, and CPC).

However, The spatial pattern of the GWS derived from GRACE is actually significantly dependent on the choice of hydrological models. Here I simply show the GWS results using individual models instead of the average. Figure 3.24 shows the derived GWS figures based on six hydrological models. The first four figures from the top two panels show the results from the individual models (NOAH, VIC, MOSAIC, and CPC) selected by *Feng et al.* [2013], and the last two figures from the bottom panel show the results from two additional models, including CLM (also from GLDAS but neglected in *Feng et al.* [2013]) and ERA (from ECMWF ERA-interim). Clearly, GWS using different hydrological models show various spatial patterns, and especially the last two models, ERA and CLM, shift the negative dome far away from the NCP area. Therefore, Figure 3.24 shows that hydrological model is still the main error source for the GWS estimation from GRACE.

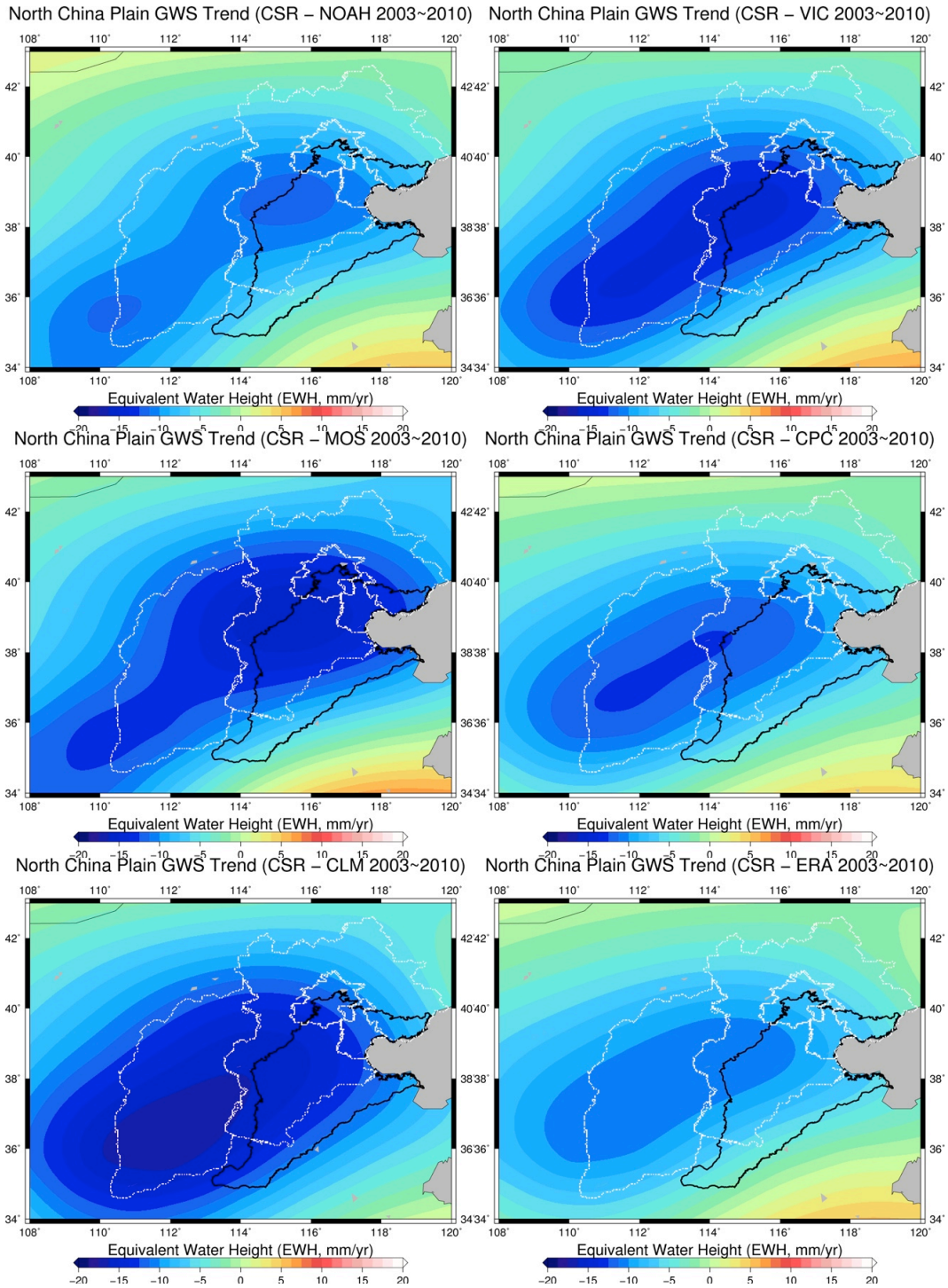


Figure 3.24 Ground Water Storage (GWS) trend map from 2003 to 2010 over North China Plain based on individual hydrological models: NOAH, VIC, MOSAIC, CPC, CLM, ERA from top left to bottom right.

3.3.2.2 Sub-monthly Terrestrial Water Storage from GRACE

If one were only interested in the inter-annual signal or annual signal, monthly sampled solutions are sufficient to estimate inter-annual signals or annual signals. But when one need to examine more rapid change signals, such as sub-monthly signal, the monthly solutions then can hardly fulfill the purpose. In this part, the 11-day solutions with 1-day step are used to study the sub-monthly variation of the TWS over NCP from GRACE. Six months' geopotential difference data from July 2003 through December 2003 are used to generate 184 11-day solutions with 1-day step up to degree and order 60. Both 200 km Gaussian smoothing and de-correlation are applied to those 11-day solutions for post-processing.

Figure 3.26 exhibits the sub-monthly analysis over North China Plain in terms of TWS as time series from July 2003 to December 2003. In Figure 3.26a the monthly TWS variation are displayed over NCP, estimated from both CSR RL05 product and my product, called OSU monthly here. The two time series can fit very well with each other, and of course each time series only contain 6 points during that period by using monthly product. Next, an additional line is added in Figure 3.26b to show the time series of the sub-monthly TWS estimated from 184 11-day solutions. The referenced epoch for both Figure 3.26a and 3.26b is the first day during this period, i.e. July 1st, 2003. For the 11-day solutions, the value from referenced epoch is subtracted from all the other solutions, and for the monthly solution, the value of the referenced epoch is interpolated and subtracted from monthly solutions. Compared to the monthly solutions, the 11-day solutions with 1-day step can reveal many high frequency features besides the low-frequency seasonal signal based on the time series. For example, from August to September 2003, the time series of the 11-day solutions deceases first then increases, indicating there might exist a short-time water loss event that is not observed by using monthly product; also for the October 2003, it exposes that there seems to exist a rapid water gain and loss within that month.

The next question is that whether or not those high-frequency feathers represent real high-frequency hydrological signal instead of purely measurement noise. Here, hydrological models are used to assess those sub-monthly solutions from GRACE. Unlike previous monthly analysis, not all the hydrological models provide sub-monthly product. Besides, since the purpose here is to examine the high-frequency fluctuation of the TWS, using the average hydrology model as the previous studies [*Rodell et al.*, 2009; *Feng et al.*, 2013] might diminish the short-time change existed in different models. Therefore, instead of compute a mean hydrology model, here two individual hydrological models are used and compared separately with GRACE observation. The first hydrological model is from GLDAS-2 product. The second one is from ERA-interim product. The components of soil moisture and snow depth from both models are used to simulate the TWS, and both of them are sub-daily sampled grid data. In order to make them consistent with GRACE observation, each grid data are transformed into spherical harmonics, truncate it to degree 60 and apply the same 200 km Gaussian smoothing. Then for each day the 11-day average are calculated from those sub-daily processed

models to compare with the 11-day solutions from GRACE. The referenced epoch is also set to July 1st. In Figure 3.26c and 3.26d, the time series of TWS change from both models are plotted, together with TWS change from GRACE 11-day solutions. Both models show similar increasing trend over these 6 months, but apparently ERA-interim agrees with GRACE solutions better than GLDAS-2. It is interesting to notice that ERA-interim model also shows two similar rapid change events in both August and October, implying that it might be more consistent with GRACE compared to GLDAS-2. The reason could be that GRACE use AOD1B data to remove the high-frequency variation from both atmosphere and ocean, and the input of the atmosphere part is the atmosphere field from ECMWF operational analysis data, which should be consistent with the soil moisture and snow depth field from ERA-Interim (i.e., ECMWF Re-Analysis).

3.3.2.3 Sub-monthly Ground Water Storage from GRACE

Since there still exists large discrepancy between the two lines in both Figure 3.26c and Figure 3.26d, it is thus obvious that soil moisture and snow depth components from models cannot fully explain the TWS change observed from GRACE. Another major contribution to the TWS change over North China Plain is the groundwater. In fact, since North China Plain includes one shallow unconfined aquifer and three deep confined aquifers, large amount of the groundwater from these aquifers has been used for agricultural irrigation, the excessive pumping of groundwater has caused severe groundwater depletion. Several studies based on GRACE monthly solution have already confirmed this depletion, and the estimated groundwater loss rate varies from 1.1–2.4 cm/yr (e.g., *Zhong et al.* [2009]; *Moiwo et al.* [2009]; *Su et al.* [2011]; *Feng et al.* [2013]).

Therefore, it is necessary to include the component from Ground Water Storage (GWS), which comes from the water table data measured from in-situ monitoring wells (courtesy of Wei Feng, IGG, CAS). The geo-locations of the 40 in-situ monitoring well stations are shown as the black points in Figure 3.25. These well stations provided daily water table depth measurements from 2002 to 2012. All the time series of the water table depth from these stations are directly plotted in Figure 3.26e, which shows very large biases between data time series from these well stations because they are located in different depth. In Figure 3.26f, the individual biases are removed to compute the mean time series as the red line. A mean specific yield number of 0.06 [*Feng et al.*, 2013] are adopted to convert the water table depth to groundwater.

If assuming the average time series can represent the groundwater change over North China Plain, then the sum of groundwater change (from average of the in-situ data) and soil moisture and snow depth components (from models) should match the GRACE observed TWS change. In Figure 3.26g and 3.26h, the groundwater change (green line) from in-situ data is added to soil moisture change from two different models (blue line) to get the sum of them (black line). As it can be seen, adding groundwater does mitigate part of the discrepancy with the GRACE, but still fails to explain some of the high frequency features. On the other hand, because of the large different between the ground water data

(Figure 3.26e), the use of mean value to represent the ground water change over a large area is also questionable.

Therefore, instead of using the average of the ground water data from all the well stations, each station is evaluated individually. In order to do that, the soil moisture and snow depth from ERA-Interim model are adopted, and subtracted from GRACE TWS, to get the GRACE derived sub-monthly ground water storage (GWS) change. Then the GRACE derived GWS are compared with individual GWS measured from each well station. As a result, 7 stations show highly correlated in-situ measurement with GRACE GWS observation. In Figure 3.27, the 7 individual GWS time series are plotted with respect to the GRACE derived GWS from 3.27b to 3.27h. Figure 3.27a shows the average of the ground water time series for comparison, which shows the average of the ground water data from all 40 well stations does moderate the high frequency feathers hidden in individual station. The time series from 3.27b to 3.27h all show similar ground water depletion near the end of August. Two time series (from station 6 and station 7) also show similar deletion in December. Figure 3.25 highlights the geo-location of the 7 well stations as the red dot, which shows that all the 7 stations are located in a small area of NCP, and also station 6 and 7 are very close to each other and therefore they show similar ground water measurement. This small area is near a shallow aquifer and also close to the high irrigation region, so the rapid change of ground water storage could be due to the short time heavy irrigation, which might possibly explain the correlation between the in-situ data and GRACE results.

The correlation between models, in-situ GWS and GRACE derived TWS and GWS implies the possibility of using GRACE to monitor the rapid change of both surface water and ground water storage change, which may eventually benefit the potential water resources management as well as improve our understanding of hydrologic circulation and discharge.

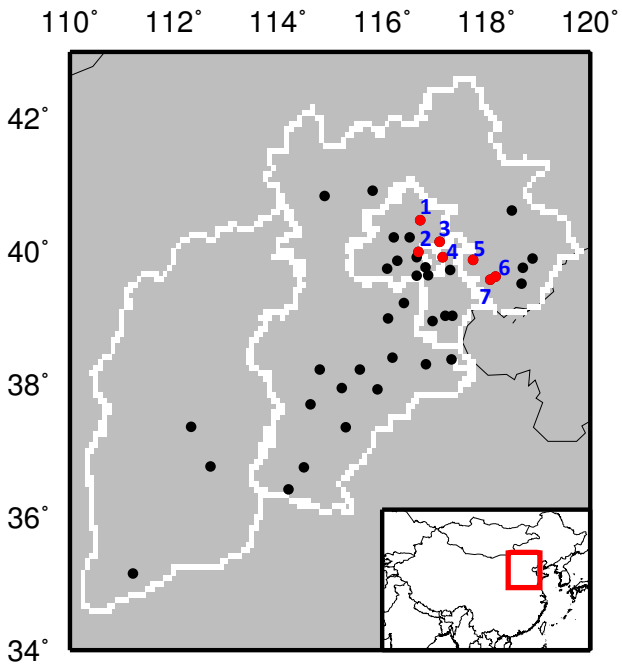


Figure 3.25 Groundwater monitoring well distribution over North China Plain. Black dot: geo-location of 40 well stations. Red dot: the 7 stations show the in-situ data which are highly correlated with GRACE derived groundwater storage change.

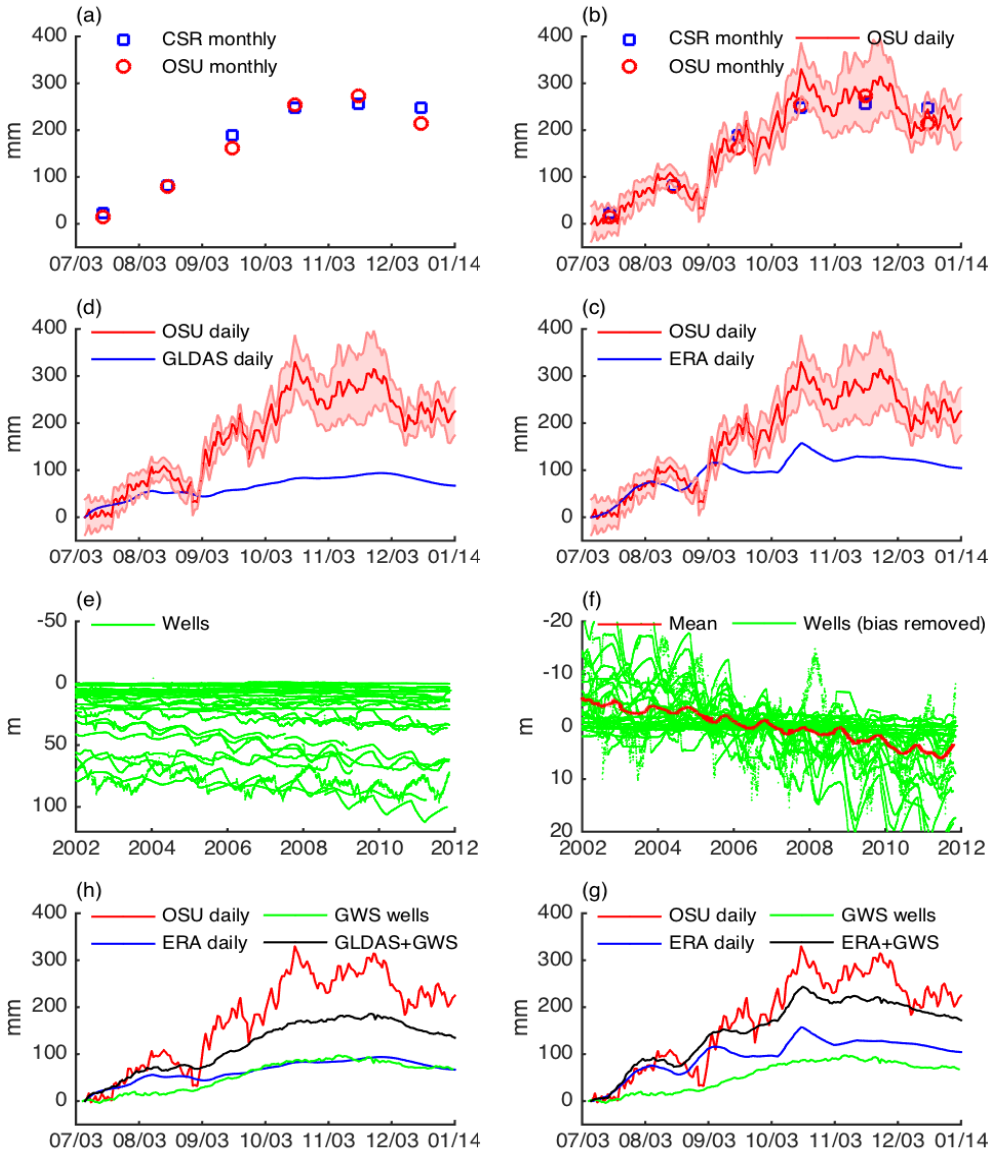


Figure 3.26 Sub-monthly GRACE Terrestrial Water Storage (TWS) over North China Plain from July to December 2003. (a) Monthly TWS from CSR RL05 and this study. (b) Daily TWS from this study. (c) Comparison of daily TWS from GRACE and model from ERA-interim. (d) Comparison of daily TWS from GRACE and model from GLDAS. (e) Water table data from 40 in-situ well stations. (f) Average of all the water table data. (g) Comparison of GRACE TWS with in-situ ground water data and model from ERA-interim. (h) Comparison of GRACE TWS with in-situ ground water data and model from GLDAS.

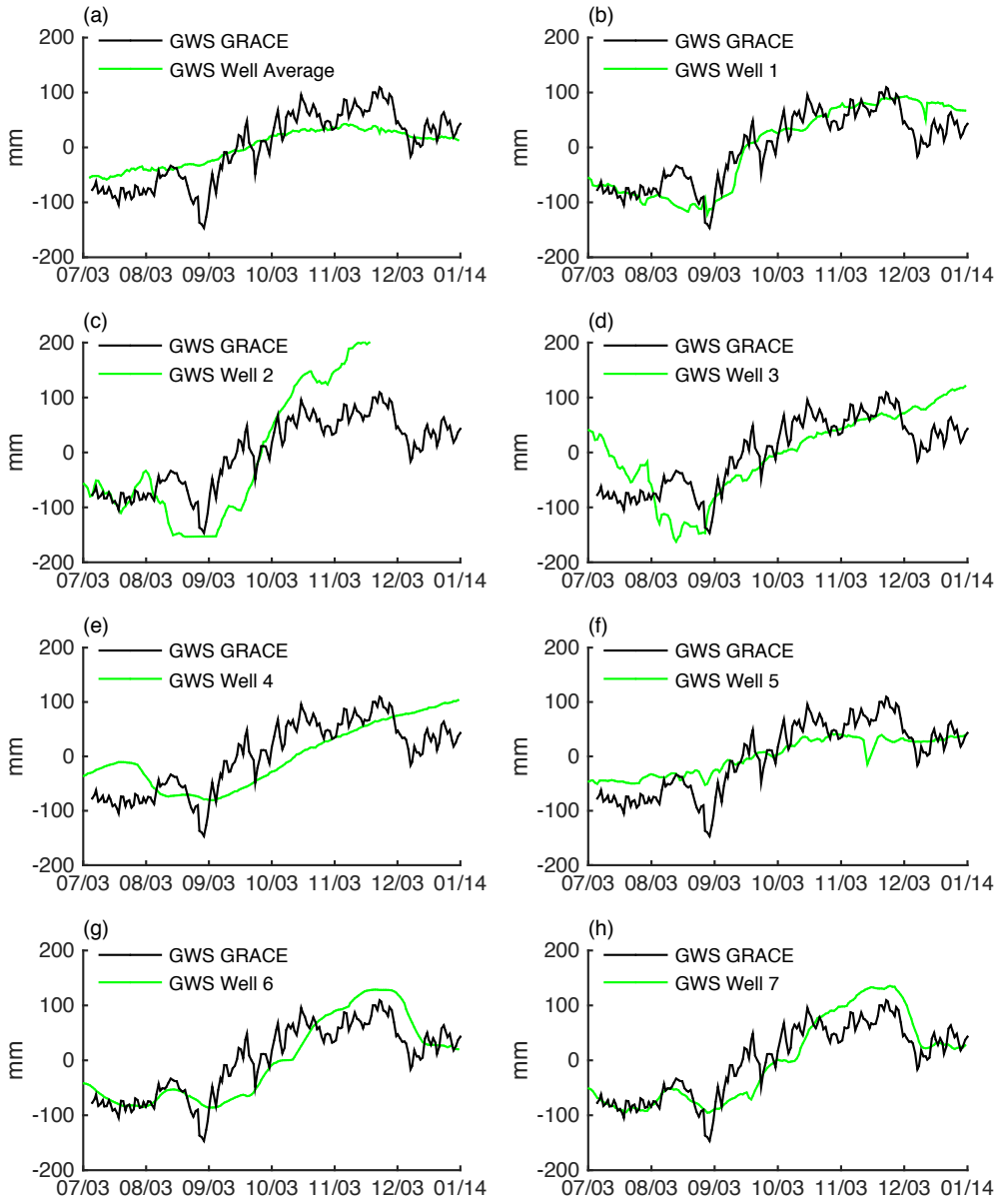


Figure 3.27 Sub-monthly GRACE Ground Water Storage (GWS) over North China Plain from July to December 2003. (a) Comparison of GRACE GWS and average of the in-situ ground water data. (b~h) Comparison of GRACE GWS and the individual in-situ ground water data from 7 different stations.

3.3.3 Measuring Snow and Ice Storm from GRACE

3.3.3.1 2008 Southeast China Snow and Ice Storm

Extreme weather events significantly impact human well being. Some of these unexpected natural hazards have led to significant loss of human lives and economics.. The 2008 Southeast China snow and ice storm [*Stone*, 2008], lasting from middle January to middle February, affected 21 out of China's 34 provinces and regions with heavy snows, ice and freezing rains. As China's worst winter in 5 decades, the storm caused extensive damage and transportation disruption, displaced nearly 1.7 million people, and claimed 129 lives. The massive accumulation of snow and ice due to the month-long precipitation and below freezing temperature even slightly changed the gravity and mass balance on Earth surface, and has consequently perturbed the relative motion of the two GRACE satellites, which is captured by KBR range-rate measurements.

Previous applications of GRACE mission are mainly limited to secular and periodic (seasonal) gravity or mass changes over a large-scaled basin (such as Amazon) and ice sheets (such as Greenland or Antarctic), or a sudden permanent gravity jump with respect to the secular and periodic change (such as earthquake), using a series of gravity solutions approximately representing the mean gravity during each calendar month. However, this traditional method cannot fulfill the purpose of this study because the cumulated gravity change caused by snow and ice storm during such a short time has neither secular/periodic change nor permanent jump of gravity. The expecting gravity change over a certain area during the storm event should increase first because of snow accumulation and decrease afterwards because of melting and runoff.

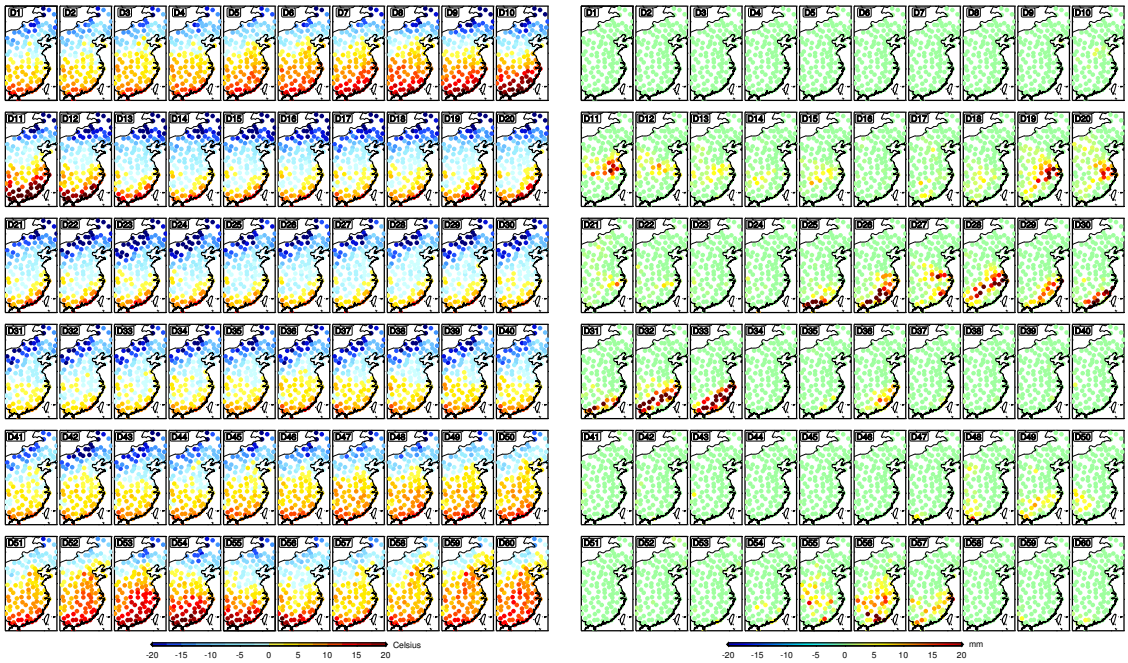


Figure 3.28 In-situ data is from 194 regional climatology stations in Mainland China (http://cdc.cma.gov.cn) from January 1st to February 29th, 2008. Left: daily mean air temperature. Right: daily precipitation.

Therefore, the new 11-day solutions with 1-day step have to be used in order to study this event. The chosen time span is exactly 60 days from January 1st to February 29th, 2008. First, the daily in-situ precipitation and temperature are shown in Figure 3.28, where the data are collected from 194 regional climatology stations over Mainland China (http://cdc.cma.gov.cn). The number for each figure is defined as the days since January 1st, 2008.

As can be seen in Figure 3.28, this severe winter hazard started at the beginning of 2008. Around the start of 2008, a warm, moist maritime tropical (mT) air mass was driven from Bay of Bengal and South China Sea to southern and central China, which abnormally increases the surface temperature to over 20 °C (D1~D10 in Figure 3.28 left) and meanwhile brings plenty amount of moisture. Around January 10th, 2008, this moist air mass was hit by a large dome of cold, continental polar (cP) air mass coming from central Asia and Mongolia. The encounter of the two air masses not only just rapidly drop the temperate below freezing point (after D11 in Figure 3.28 left), but more seriously suddenly develop into a stable atmospheric boundary layer, which is a favorable configuration for continuous snow and freezing rains over a large region [*Zhou et al.*, 2011].

As a result, snow and freezing rain started around January 11th, which can be seen from the precipitation maps on the right of Figure 3.28. The snow struck the east-central China

first, mainly over middle and lower reaches of the Yangtze River, including provinces of Anhui, Hunan, Henan, Shandong, Jiangsu Provinces and municipality of Shanghai. After that, around the second half of January, the affected area expanded into larger southeast China driven by the slow moving of the cold cP air mass towards south, which eventually bring the record-breaking snow, ice and freezing rains to half of China.

From the right panel of Figure 3.28, the successive four separate waves (10–16 January, 18–22 January, 25– 29 January, and 31 January–6 February) of snow and ice snow last until early February, and finally stop around February 6th (D37). During that period, almost all the recorded mean air temperature around southeast China is below or around 0 °C (Figure 3.28 left). Therefore the snow and ice could continuously accumulate onsite instead of evapotranspiration and runoff. The short interval between the snow, such as January 23rd and 24th, could only temporally slow down the accumulation, but not long enough for ice melt also because of the low temperature.

3.3.3.2 Results from GRACE

The gravity change due to the surface snow and ice change can be measured by GRACE. To see that, the geopotential differences during that two months are collected, including 5 more days for each end, and the strategy of 11-day solutions with 1-day step are applied again to solve the global gravity field up to degree and order 60. After applying de-correlation and 200 km Gaussian smoothing, the EWH are computed for each 11-day solution. The referenced field is chosen as January 1st, 2008, which is computed using the data between December 27th, 2007 and January 6th, 2008. During that period, there is no evident precipitation for the area so the gravity fluctuation between December 27th, 2007 and January 6th, 2008 is assumed to be negligible. Therefore, the results actually represent the 11-day average of gravity change for each day with respect to the reference day, which is mainly caused by the snow accumulation and melting afterwards. The gravity changes in terms of EWH during those two months are shows in Figure 3.29.

Figure 3.29 clearly reveals the process of surface mass increasing and deceasing caused by the accumulation and melting of the snow and ice during this winter event. Surface mass first accumulate mainly over east-central China, since that is where the cold air initially struck. The accumulation in the east reached the peak around January 22, right before the short two days interval. After that, the surface mass accumulation over the southeast China overwhelmed the east, and reach to the maximum on February 6th (D37) when the major precipitation stopped. With the temperature gradually increasing after February 6th, the accumulated snow began to melt, so the evapotranspiration and runoff dominated the snow change and decreased the surface mass.

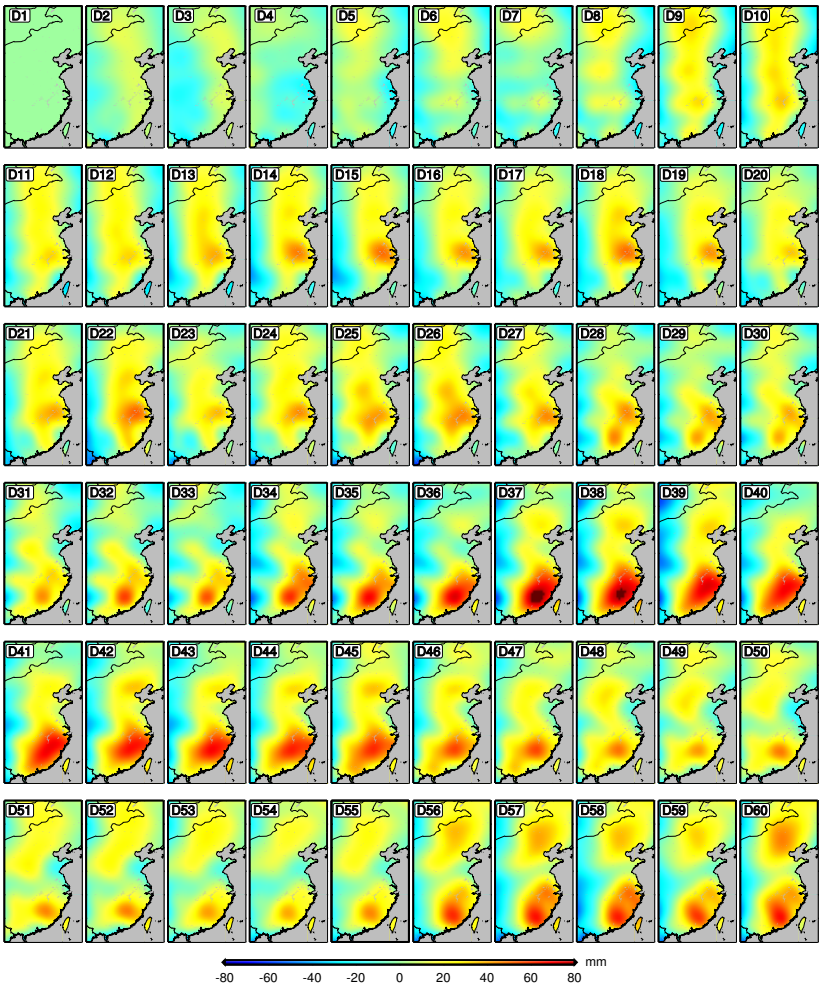


Figure 3.29 GRACE derived EWH from January 1st to February 29th, 2008, which captured the surface mass change signal caused by 2008 Southeast China snow and ice storm.

3.3.3.3 Validation using Model and In-situ Data

In order to further validate the GRACE-derived sub-monthly solution, the model from ERA-Interim product are used to compare with the EWH from GRACE. ERA-Interim is a global atmospheric reanalysis product, including global atmospheric and surface parameters from 1 January 1979 to present, which is based on ECMWF global daily gridded model of various hydrology parameters, such as precipitation (P), evapotranspiration (ET) and runoff (R). Using these parameters, the daily EWH can be modeled as

$$EWH = \int_{t_0}^{t} P + ET - R\,, \tag{3.6}$$

where the integral starts on January 1st until each day. The resulting EWH model is the gridded data, which have to be transformed into spherical harmonics and truncated to degree and order 60. Then the 11-day moving average and the 200 km Gaussian smoothing are applied, in order to be consistent with the GRACE-derived EWH. Besides, the daily in-situ evaporation observations from the 194 regional climatology stations are combined with the precipitation observations to model the EWH using the same method. Here the runoff term is ignored.

Figure 3.30 shows the results of the EWH modeling from ERA-Interim and in-situ data. Compared to GRACE result, model from ERA-Interim shows smaller amplitude, and model from in-situ data shows larger amplitude, which is mainly caused by the lack of runoff in-situ data. The GRACE time series is found to be closer with the evolution of the in-situ data. But both the model and the in-situ data show similar EWH change pattern as GRACE result, i.e., accumulation in the east first, then expand to southeast and then fade away. This similarity again confirms my GRACE 11-day solutions with 1-day step do capture the surface mass change signal caused by 2008 Southeast China snow and ice storm, and furthermore prove that GRACE can provide an additional constraint to understand the mass transportation and the dynamic processes during extreme and rapid changing weather events, especially for the poor gauged area.

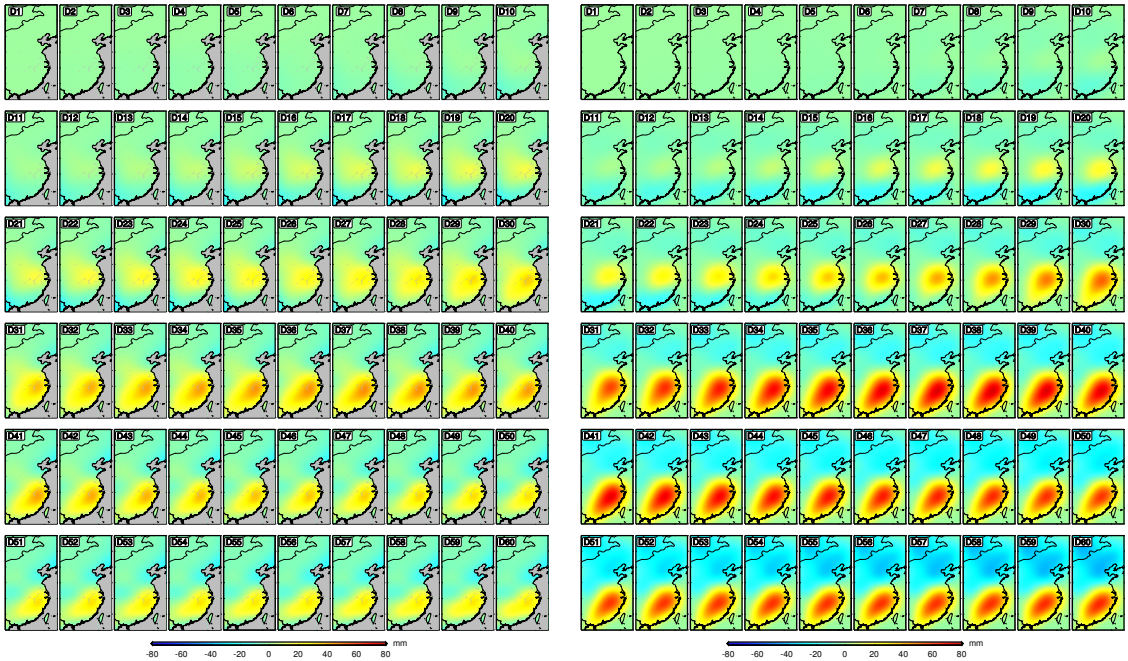


Figure 3.30 GRACE-type EWH modeling using ERA-Interim (left) and in-situ data (right), from January 1st to February 29th, 2008.

Chapter 4 Conclusions

A *new* approach, called improved energy balance formalism, has been developed for time-variable gravity field recovery from GRACE satellites in this study. A simple, innovative method, called the alignment equation, is first presented to exactly utilize the most accurate measurements, i.e., range-rate, from the satellite-to-satellite tracking system. Next, a reconstruction algorithm is also used to overcome the possible systematic error caused by the coupling between reference orbit and gravity, and recalibrate the accelerometer by uncovering an offset of scale parameter after 2010. Finally, a more precise formulation of energy equation to calculate the geopotential difference observables is adopted by considering more rigorous modeling and validated by using sophisticated close-loop simulation. Compared to the previous study about energy balance approach, this improved approach can better preserve gravity information from range-rate data and reduce error from orbit data, which would essentially improve the estimation of geopotential difference. Analysis of more than 10 years of GRACE data indicated that the resulting geopotential difference estimates agree well with predicted values from official GRACE RL05 monthly solutions: with much higher correlation at 0.9, as compared to 0.5–0.8 reported by previous energy balance studies. The developed formalism is applicable to the general case of low-low satellite-to-satellite radiometric or laser interferometric tracking measurements, such as GRACE Follow-on or other Next Generation Gravity Field missions, for efficient retrieval and studies of Earth's mass transport evolutions.

The *first* GRACE global monthly solution series based on energy conservation principle is produced from 2003 to 2013 using the geopotential difference data and a straightforward inversion method up to degree and order 60. The recovered monthly solutions show comparable signal-to-noise ratio with official GRACE RL05 monthly solutions in both spatial domain and spectral domain, for both the month with homogeneous ground track distribution and the month with non-homogeneous ground track distribution due to the near-repeat orbit. Further comparison also indicated consistent secular and seasonal gravity variation, for both global trend and regional basins, with and without any post-processing techniques.

This study demonstrates that an *enhanced* temporal resolution is achievable using the improved energy balance method because of the flexible inversion process from geopotential difference data. The enhanced solutions are conducted over Greenland first and reveal that a substantially higher temporal resolution is achievable at 10 or 11-day interval from GRACE data, as compared to the official monthly solutions, but without the compromise of spatial resolution, nor the need to use regularization or post-processing.

The similar strategy is further applied to solve gravity field using the neighboring 11-day geopotential difference data but with only 1-day step, leading to *sub-monthly* solutions for the purpose of studying the gravity variation within a month. These 11-day solutions with 1-day step have been applied to study both the terrestrial and ground water storage over North China Plain aquifer. Analysis of 6 months' sub-monthly solutions show the GRACE-derived high-frequency variation of TWS is more consistent with the models from ERA-Interim than GLDAS-2. After removing the soil moisture using ERA-Interim, correlation between the derived GWS and selected in-situ measurements from well stations can be identified, which implies the possibility of using GRACE to monitor the rapid change of both surface water and ground water storage change, and could eventually benefit the potential water resources management as well as improve our understanding of hydrologic circulation and discharge.

Finally, the 11-day solutions with 1-day step are applied to study the mass variations during the event of 2008 Southeast China snow and ice storm from January to February 2008. The surface mass change caused by the rapid snow and ice accumulation and melting has been successfully captured by GRACE solutions after routine post-processing, which are also validated by both the in-situ data and hydrologic models, and demonstrates that these 11-day solutions can provide an additional constraint to understand the mass transportation and the dynamic processes during extreme and rapid changing weather events, especially for the poor gauged area.

Several limitations still exist in the current realization of the improved energy balance formalism. The fore problem is that signal-coherent adjustment is still inevitable, which means that the geopotential differences, gravity solutions, reconstructed orbit and the solved systematic parameters are all biased to the *a priori* gravity model. Iteration might be able to overcome this problem and lead to a more independent monthly global solutions series. Another problem is the purely dynamic orbit based on the orbit reconstruction algorithm is, although efficient, but still a simplified method, which may suffer the unexpected error from the input orbit. In the future, the reconstruction algorithm should be replaced by directly solving dynamic orbit from GPS tracking data. The third imperfection is the formulation of energy method still contains minor approximation on the order of 10^{-5} m^2/s^2, which is however negligible definitely for processing GRACE and also most likely for the future GRACE Follow-on observations.

Therefore, the new method from this study would also benefit the forthcoming GRACE follow-on mission, especially considering the possibility that the precision of range-rate data can be improved by up to a factor of 20 [*Loomis et al.*, 2012], but the precision of GPS tracking data may not have significant advances. In that case, for the traditional conventional dynamic method the weighting of GPS tracking data would need to be further reduced, which could be more analogous to this method since GPS data and range-rate data have already separately handled through the alignment equation. Therefore, the energy balance method might have a unique contribution to the processing of more accurate data from next generation satellite gravimetry mission to extend Earth's mass transport climate record.

Bibliography

Badura, T., C. Sakulin, C. Gruber, and R. Klostius (2006), Derivation of the CHAMP-only global gravity field model TUG-CHAMP04 applying the energy integral approach, *Studia Geophysica et Geodaetica*, 50(1), 59-74, doi: 10.1007/s11200-006-0002-3.

Bettadpur, S. (2009), Recommendation for a-priori bias & scale parameters for level-1B ACC data (Version 2). GRACE TN-02.

Bettadpur, S. (2012), UTCSR Level-2 processing standards document for Level-2 product release 0005, GRACE 327-742, Rev. 4.0, Center for Space Research, The University of Texas at Austin.

Biancale, R., and A. Bode (2006), Mean annual and seasonal atmospheric tide models based on 3-hourly and 6-hourly ECMWF surface pressure data, *Tech. Rep.*, Potsdam, Deutsches GeoForschungsZentrum GFZ, doi: http://doi.org/10.2312/GFZ.b103-06011.

Bjerhammar, A. (1969), On the energy integral for satellites. *Tellus*, 21: 1–9. doi: 10.1111/j.2153-3490.1969.tb00412.x

Bruinsma, S., J. M. Lemoine, R. Biancale, and N. Vales (2010), CNES/GRGS 10-day gravity field models (release 2) and their evaluation, *Advances in Space Research*, 45(4), 587–601, doi:10.1016/j.asr.2009.10.012.

Case, K., G. Kruizinga, S.C. Wu (2010), GRACE level 1B data product user handbook, NASA Jet Propulsion Laboratory, revision 1.3, JPL D-22027, GRACE 327-733.

Cazenave, A., and J. Chen (2010), Time-variable gravity from space and present-day mass redistribution in the Earth system. *Earth Planet. Sci. Lett.*, 298(3), 263-274, doi:10.1016/j.epsl.2010.07.035.

Chambers, DP (2006), Evaluation of new GRACE time-variable gravity data over the ocean, *Geophys Res Lett* 33 (17):L17603. doi:10.1029/2006gl027296

Chen, J.L., C.R. Wilson, and T.B. Tapley (2006) Satellite gravity measurements confirm accelerated melting of Greenland ice sheet, *Science* 313 (5795):1958-1960. doi:10.1126/science.1129007

Chen, J. L., C.R. Wilson, and B.D. Tapley (2010), The 2009 exceptional Amazon flood and interannual terrestrial water storage change observed by GRACE, *Water Resources Research*, 46:12526, doi:10.1029/2010WR009383.

Chen, J.L., C.R. Wilson, B.D. Tapley, and S. Grand (2007), GRACE detects coseismic and postseismic deformation from the Sumatra-Andaman earthquake, *Geophys Res Lett* 34 (13):L13302. doi:10.1029/2007gl030356

Chen, J. L., C.R. Wilson, and B.D. Tapley, Z.L. Yang, and G.Y. Niu (2009), 2005 drought event in the Amazon River basin as measured by GRACE and estimated by climate models, *J. Geophys. Res.*, 114(B13):5404, doi:10.1029/2008JB006056.

Chen, Y. Q., B. Schaffrin, and C. K. Shum (2008), Continental water storage changes from GRACE Line-of-sight range acceleration measurements, in *VI Hotine-Marussi Symposium on Theoretical and Computational Geodesy*, Peiliang Xu, Jingnan Liu, Athanasios Dermanis (Ed.), International Association of Geodesy Symposia, Volume 132, ISSN 0939-9585, pp 62-66, doi: 10.1007978-3-540-74584-6.

Cheng, M., B. D. Tapley, and J. C. Ries (2013), Deceleration in the Earth's oblateness, J. Geophys. Res. Solid Earth, 118, 740–747, doi:10.1002/jgrb.50058.

Dahle, C., F. Flechtner, C. Gruber, D. König, R. König, G. Michalak, and K.-H. Neumayer (2012), GFZ GRACE Level-2 Processing Standards Document for Level-2 Product Release 0005, *Tech. Rep.*, Potsdam, Deutsches GeoForschungsZentrum GFZ, doi: 10.2312/GFZ.b103-12020.

Dai, C., C.K. Shum, R. Wang, L. Wang, J. Guo, K. Shang, and B. Tapley (2014), Improved constraints on seismic source parameters of the 2011 Tohoku earthquake from GRACE gravity and gravity gradient changes, *Geophysical Research Letters*, 41:1929–1936, doi:10.1002/2013GL059178.

Davis, J.L., M.E. Tamisiea, P. Elosegui, J.X. Mitrovica, and E.M. Hill (2008), A statistical filtering approach for Gravity Recovery and Climate Experiment (GRACE) gravity data, *J Geophys Res* 113 (B4):B04410. doi:10.1029/2007jb005043

de Paiva, R. C. D., D.C. Buarque, W. Collischonn, M.P. Bonnet, F. Frappart, S. Calmant, and C.A. Bulhões Mendes (2013), Large-scale hydrologic and hydrodynamic modeling of the Amazon River basin, *Water Resources Research*, 49:1226–1243, doi:10.1002/wrcr.20067.

Ditmar, P, J. Teixeira da Encarnação, and H. Hashemi Farahani (2012), Understanding data noise in gravity field recovery on the basis of inter-satellite ranging measurements acquired by the satellite gravimetry mission GRACE, *J. Geod.*, 86:441–465, doi:10.1007/s00190-011-0531-6.

Döll, P., H. Müller Schmied, C. Schuh, F. Y. Portmann, and A. Eicker (2014), Global-scale assessment of groundwater depletion and related groundwater abstractions: Combining hydrological modeling with information from well observations and GRACE satellites, *Water Resources Research*, 50:5698–5720, doi:10.1002/2014WR015595.

Duan, X.J., J.Y. Guo, C.K. Shum, and W. van der Wal (2009), On the postprocessing removal of correlated errors in GRACE temporal gravity field solutions, *J Geod* 83 (11):1095-1106. doi:10.1007/s00190-009-0327-0

Feng, W., M. Zhong, J.-M. Lemoine, R. Biancale, H.-T. Hsu, and J. Xia (2013), Evaluation of groundwater depletion in North China using the Gravity Recovery and Climate Experiment (GRACE) data and ground-based measurements, *Water Resources Research*, 49:2110–2118, doi:10.1002/wrcr.20192.

Gerlach, Ch., L. Földvary, D. Švehla, Th. Gruber, M. Wermuth, N. Sneeuw, B. Frommknecht, H. Oberndorfer, Th. Peters, M. Rothacher, R. Rummel and P. Steigenberger (2003), A CHAMP-only gravity field model from kinematic orbits using the energy integral, *Geophys. Res. Lett.*, 30(20), 2037, doi:10.1029/2003GL018025.

Guo, J. Y., X. J. Duan, and C.K. Shum (2010), Non-isotropic Gaussian smoothing and leakage reduction for determining mass changes over land and ocean using GRACE data. *Geophys. J. Int.*, 181: 290–302, doi: 10.1111/j.1365-246X.2010.04534.x.

Guo, J.Y., K. Shang, C. Jekeli, and C.K. Shum, On the energy integral formulation of gravitational potential differences from satellite-to- satellite tracking, *Celestial Mech. Dyn. Astron.*, 121, Issue 4, 415–429, doi:10.1007/s10569-015-9610-y, 2015.

Han, S.-C., C. Jekeli, and C.K. Shum (2002), Efficient gravity field recovery using in-situ disturbing potential observables from CHAMP, *Geophys. Res. Lett.*, 29(16), 1789, doi:10.1029/2002GL015180.

Han, S.-C. (2003) Efficient global gravity determination from satellite-to-satellite tracking (SST). *Diss. PhD thesis*, Geodetic and Geoinformation Science, Department of Civil and Environmental Engineering and Geodetic Science, The Ohio State University, Columbus.

Han, S.-C., H. Kim, I.-Y. Yeo, P. Yeh, T. Oki, K.-W. Seo, D. Alsdorf, and S. B. Luthcke (2009), Dynamics of surface water storage in the Amazon inferred from measurements of inter-satellite distance change, *Geophys. Res. Lett.*, 36, L09403, doi:10.1029/2009GL037910.

Han, S.-C., C.K. Shum, M. Bevis, C. Ji, and C.-Y. Kuo (2006), Crustal Dilatation Observed by GRACE After the 2004 Sumatra-Andaman Earthquake, *Science*, 313:658–662, doi:10.1126/science.1128661.

Han, S.-C., C.K. Shum, and C. Jekeli (2006), Precise estimation of in-situ geopotential differences from GRACE low-low satellite-to-satellite tracking and accelerometer data, *J. Geophys. Res.*, 111, B04411, doi:10.1029/ 2005JB003719.

Han, S.-C., C.K. Shum, C. Jekeli, and D. Alsdorf (2005a), Improved estimation of terrestrial water storage changes from GRACE, *Geophys. Res. Lett.*, 32, L07302, doi:10.1029/2005GL022382.

Han, S.-C., C.K. Shum, and C. Jekeli, C.Y. Kuo, C. Wilson, and K.W. Seo (2005b), Non-isotropic filtering of GRACE temporal gravity for geophysical signal enhancement. *Geophys J Int*, 163 (1):18-25. doi:10.1111/j.1365-246X.2005.02756.x

Huang, Z., Y. Pan, H. Gong, P. J. Yeh, X. Li, D. Zhou, and W. Zhao (2015), Subregional-scale groundwater depletion detected by GRACE for both shallow and deep aquifers in North China Plain. *Geophys. Res. Lett.*, 42, 1791–1799. doi: 10.1002/2014GL062498.

Jäggi, A., H. Bock, R. Pail, and H. Goiginger (2008), Highly-reduced dynamic orbits and their use for global gravity field recovery: a simulation study for GOCE, *Studia Geophysica et Geodaetica*, 52(3), 341-359, doi: 10.1007/s11200-008-0025-z.

Jekeli, C. (1999), The determination of gravitational potential differences from satellite-to-satellite tracking, *Celestial Mech. Dyn. Astron.*, 75, 85-100, doi: 10.1023/A:1008313405488.

Jekeli, C. (2001), *Inertial navigation systems with geodetic applications*, Walter de Gruyter, New York.

Johnson, G. C., and D. P. Chambers (2013), Ocean bottom pressure seasonal cycles and decadal trends from GRACE Release-05: Ocean circulation implications, *Journal of Geophysical Research (Oceans)*, 118:4228–4240, doi:10.1002/jgrc.20307.

Kang, Z., B. Tapley, S. Bettadpur, J. Ries, and P. Nagel (2006a), Precise orbit determination for GRACE using accelerometer data. *Adv. Space Res.*, 38(9), 2131-2136.

Kang, Z., B. Tapley, S. Bettadpur, J. Ries, P. Nagel, and R. Pastor (2006b), Precise orbit determination for the GRACE mission using only GPS data, *J. Geod.*, 80(6), 322-331.

Kang, Z., S. Bettadpur, P. Nagel, R. Pastor, T. Pekker, S. Poole, and B. Tapley (2008), Quick-Look Gravity Solutions From GRACE, In AGU Fall Meeting Abstracts.

Klees, R, E.A. Revtova, B.C. Gunter, P. Ditmar, E. Oudman, H.C. Winsemius, and H.H.G. Savenije (2008), The design of an optimal filter for monthly GRACE gravity models, *Geophys J Int*, 175 (2):417-432. doi:10.1111/j.1365-246X.2008.03922.x

König, D., and C. Dahle (2014), GRACE Gravity Modeling Using the Integrated Approach, *In Observation of the System Earth from Space-CHAMP, GRACE, GOCE and future missions* (pp. 41-45). Springer Berlin Heidelberg.

Konopliv, A. S., et al. (2013), The JPL lunar gravity field to spherical harmonic degree 660 from the GRAIL Primary Mission, *J. Geophys. Res. Planets*, 118, 1415–1434, doi:10.1002/jgre.20097.

Kurtenbach, E., T. Mayer-Gürr, and A. Eicker (2009), Deriving daily snapshots of the Earth's gravity field from GRACE L1B data using Kalman filtering, *Geophys. Res. Lett.*, 36, L17102, doi:10.1029/2009GL039564.

Kurtenbach, E., A. Eicker, T. Mayer-Gürr, M. Holschneider, M. Hayn, M. Fuhrmann, and J. Kusche (2012), Improved daily gravity field solutions using a Kalman smoother, J. Geodyn., 59–60, doi: 10.1016/j.jog.2012.02.006

Kusche, J. (2007), Approximate decorrelation and non-isotropic smoothing of time-variable GRACE-type gravity field models, *J Geod*, 81 (11):733-749. doi:10.1007/s00190-007-0143-3

Lemoine, J.-M., S. Bruinsma, S. Loyer, R. Biancale, J.-C. Marty, F. Perosanz, and G. Balmino (2007), Temporal gravity field models inferred from GRACE data, *Adv. Space Res.*, 39, 1620–1629, doi:10.1016/j.asr.2007.03.062.

Liu, X., P. Ditmar, C. Siemes, D. C. Slobbe, E. Revtova, R. Klees, R. Riva, and Q. Zhao (2010), DEOS Mass Transport model (DMT-1) based on GRACE satellite data: methodology and validation. *Geophys. J. Int.*, 181: 769–788, doi: 10.1111/j.1365-246X.2010.04533.x.

Llovel, W., M. Becker, A. Cazenave, J. F. Crétaux, and G. Ramillien (2010), Global land water storage change from GRACE over 2002–2009; Inference on sea level, *Comptes Rendus Geoscience*, 342(3), 179-188.

Loomis, B. D., R. S. Nerem, and S. B. Luthcke (2012), Simulation study of a follow-on gravity mission to GRACE. *J. Geod.*, 86(5), 319-335, doi: 10.1007/s00190-011-0521-8.

Luthcke, S. B., D. D. Rowlands, F. G. Lemoine, S. M. Klosko, D. Chinn, and J. J. McCarthy (2006), Monthly spherical harmonic gravity field solutions determined from GRACE inter-satellite range-rate data alone, *Geophys. Res. Lett.*, 33, L02402, doi:10.1029/2005GL024846.

Matsuo, K. and K. and Heki (2012), Anomalous precipitation signatures of the Arctic Oscillation in the time-variable gravity field by GRACE, *Geophysical Journal International*, 190:1495–1506, doi:10.1111/j.1365-246X.2012.05588.x.

Mayer-Gürr, T., A. Eicker, and K. H. Ilk (2007), ITG-Grace02s: A GRACE gravity field derived from range measurements of short arcs, in *Gravity Field of the Earth: Proceedings of the 1st International Symposium of the International Gravity Field Service (IGFS)*, Special Issue 18, edited by A. Kiliçoglu and R. Forsberg, pp. 193–198, Gen. Command of Mapp., Ankara, Turkey.

Mayer-Gürr, T., R. Savcenko, W. Bosch, I. Daras, F. Flechtner, and C. Dahle (2012), Ocean tides from satellite altimetry and GRACE, *J. Geod.*, 59–60, 28–38.

Meyer, U, A. Jäggi, and G. Beutler (2012), Monthly gravity field solutions based on GRACE observations generated with the Celestial Mechanics Approach, *Earth Planet. Sci. Lett.*, 345, 72-80, doi: 10.1016/j.epsl.2012.06.026.

Morrison, F. (1970), Comments on paper by Milo Wolff, 'Direct measurements of the Earth's gravitational potential using a satellite pair', *J. Geophys. Res.*, 75(11), 2142–2143, doi:10.1029/JB075i011p02142.

Pail, R., S. Bruinsma, F. Migliaccio, C. Förste, H. Goiginger, W. D. Schuh, E. Höck, M. Reguzzoni, J. M. Brockmann, O. Abrikosov, M. Veicherts, T. Fecher, R. Mayrhofer, I. Krasbutter, F. Sansò, C. C. Tscherning (2011), First GOCE gravity field models derived by three different approaches, *J. Geod.*, 85(11), 819–843, doi:10.1007/s00190-011-0467-x.

Petit, G., and B. Luzum (2010), IERS Conventions (2010). *IERS Technical Note No. 36*. Verlag des Bundesamts für Kartographie und Geodäsie, Frankfurt am Main.

Ramillien, G., R. Biancale, S. Gratton, X. Vasseur, and S. Bourgogne (2011), GRACE-derived surface water mass anomalies by energy integral approach: Application to continental hydrology, *J. Geod.*, 6, 313–328, doi:10.1007/s00190-010-0438-7.

Ries, J. C., S. Bettadpur, S. Poole, and T. Richter (2011), Mean Background Gravity Fields for GRACE Processing, *GRACE Science Team Meeting*, Austin, TX, 8-10 August 2011.

Rodell, M., J.S. Famiglietti, J. Chen, S.I. Seneviratne, P. Viterbo, H. Holl, and C.R. Wilson (2004), Basin scale estimates of evapotranspiration using GRACE and other observations, *Geophysical Research Letters*, 31:20504–+, doi:10.1029/2004GL020873.

Rodell, M., I. Velicogna, and J. S. Famiglietti (2009), Satellite-based estimates of groundwater depletion in India, *Nature*, 460(7258), 999-1002.

Rowlands, D. D., R. D. Ray, D. S. Chinn, and F. G. Lemoine (2002), Short-arc analysis of intersatellite tracking data in a gravity mapping mission, *J. Geod.*, 76, 307–316, doi:10.1007/s00190-002-0255-8.

Rowlands, D. D., S. B. Luthcke, S. M. Klosko, F. G. Lemoine, D. S. Chinn, J. J. McCarthy, C. M. Cox, and O. B. Anderson (2005), Resolving mass flux at high spatial and temporal resolution using GRACE intersatellite measurements, *Geophys. Res. Lett.*, 32, L04310, doi:10.1029/2004GL021908.

Rowlands, D. D., S. B. Luthcke, J. J. McCarthy, S. M. Klosko, D. S. Chinn, F. G. Lemoine, J.-P. Boy, and T. J. Sabaka (2010), Global mass flux solutions from GRACE: A comparison of parameter estimation strategies—Mass concentrations versus Stokes coefficients, *J. Geophys. Res.*, 115, B01403, doi:10.1029/2009JB006546.

Save, H., S. Bettadpur, B. D. Tapley (2012), Reducing errors in the GRACE gravity solutions using regularization, *J. Geod.*, 86, 695-711, doi:10.1007/s00190-012-0548-5.

Schmidt, M., S.-C. Han, J. Kusche, L. Sanchez, and C. K. Shum (2006), Regional high-resolution spatiotemporal gravity modeling from GRACE data using spherical wavelets, *Geophys. Res. Lett.*, 33, L08403, doi:10.1029/2005GL025509.

Schmidt, M., F. Seitz, and C. K. Shum (2008), Regional four-dimensional hydrological mass variations from GRACE, atmospheric flux convergence, and river gauge data, *J. Geophys. Res.*, 113, B10402, doi:10.1029/2008JB005575.

Schrama, E.J.O., B. Wouters, and D.A. Lavallee (2007), Signal and noise in Gravity Recovery and Climate Experiment (GRACE) observed surface mass variations, *J Geophys Res*, 112 (B8): B08407. doi:10.1029/2006jb004882

Shi, X., X. Xu, and C. Lu (2010), The dynamic and thermodynamic structures associated with a series of heavy precipitation events over China during January 2008. *Weather and Forecasting*, 25(4), 1124-1141.

Simons, F., F. Dahlen, and M. Wieczorek (2006), Spatiospectral concentration on a sphere, *SIAM Rev.*, 48(3), 504–536, doi:10.1137/S0036144504445765.

Stone, R (2008), Ecologists report huge storm losses in China's forests, *Science*, 319(5868), 1318-1319.

Swenson, S., P. J.-F. Yeh, J. Wahr, and J. Famiglietti (2006a), A comparison of terrestrial water storage variations from GRACE with in-situ measurements from Illinois, *Geophys. Res. Lett.*, 33, L16401, doi:10.1029/2006GL026962.

Swenson S, J. Wahr (2006b), Post-processing removal of correlated errors in GRACE data, *Geophys Res Lett* 33 (8): L08402. doi:10.1029/2005gl025285

Tamisiea, M. E., J. X. Mitrovica, and J. L. Davis (2007), GRACE gravity data constrain ancient ice geometries and continental dynamics over Laurentia, *Science*, 316.5826: 881-883.

Tangdamrongsub, N., C. Hwang, C. K. Shum, and L. Wang (2012), Regional surface mass anomalies from GRACE KBR measurements: Application of L-curve regularization and a priori hydrological knowledge, *J. Geophys. Res.*, 117, B11406, doi:10.1029/2012JB009310.

Tapley, B. D., S. Bettadpur, M. Watkins, and C. Reigber (2004a), The gravity recovery and climate experiment: Mission overview and early results, *Geophys. Res. Lett.*, 31, L09607, doi:10.1029/2004GL019920.

Tapley, B. D., B. Schutz, and G. H. Born (2004b), *Statistical Orbit Determination*, Elsevier, New York.

van der Wal, W., P. L. Whitehouse, and E. J. O. Schrama (2015), Effect of GIA models with 3D composite mantle viscosity on GRACE mass balance estimates for Antarctica. Earth and planetary science letters.

Van Helleputte, T., E. Doornbos, and P. Visser (2009), CHAMP and GRACE accelerometer calibration by GPS-based orbit determination. *Adv. Space Res.*, 43(12), 1890-1896.

Velicogna, I. and J. Wahr (2006), Measurements of Time-Variable Gravity Show Mass Loss in Antarctica, *Science*, 311:1754–1756, doi:10.1126/science.1123785.

Visser, P.N.A.M., N. Sneeuw, and C. Gerlach (2003), Energy integral method for gravity field determination from satellite orbit coordinates, *J. Geod.*, 77, 207-216, doi: 10.1007/s00190-003-0315-8.

Wagner, C., D. McAdoo, J. Klokoçcnk, and J. Kosteleck (2006), Degradation of geopotential recovery from short repeat-cycle orbits: Application to GRACE monthly fields, *J. Geod.*, 80, 1394–1432, doi:10.1007/s00190-006-0036-x.

Wahr, J., M. Molenaar, and F. Bryan (1998), Time variability of the Earth's gravity field: Hydrological and oceanic effects and their possible detection using GRACE, *J. Geophys. Res.*, 103(B12), 30205–30229, doi:10.1029/98JB02844.

Watkins, M. M., and D. N. Yuan (2012), JPL Level-2 Processing Standards Document For Level-2 Product Release 05. Scientific Technical Report-Data, Rev. 4.0, Jet Propulsion Laboratory, Pasadena.

Wang, X., C. Gerlach, and R. Rummel (2012), Time-variable gravity field from satellite constellations using the energy integral, *Geophys. J. Int.*, 190(3), 1507-1525, doi: 10.1111/j.1365-246X.2012.05578.x.

Weigelt, M., N. Sneeuw, E. J. O. Schrama, and P. N. A. M. Visser (2013), An improved sampling rule for mapping geopotential functions of a planet from a near polar orbit, *J. Geod.*, 87(2), 127–142, doi:10.1007/s00190-012-0585-0.

Wessel, P., and W. H. F. Smith (1991), Free software helps map and display data, *Eos Trans. AGU*, 72(41), 441, doi:10.1029/90EO00319.

Wolff, M. (1969), Direct measurements of the Earth's gravitational potential using a satellite pair, *J. Geophys. Res.*, 74(22), 5295–5300, doi:10.1029/JB074i022p05295.

Wouters, B., and E.J.O. Schrama (2007), Improved accuracy of GRACE gravity solutions through empirical orthogonal function filtering of spherical harmonics, *Geophys Res Lett*, 34 (23): L23711. doi:10.1029/2007gl032098

Wu, S. C., G. Kruizinga, and W. Bertiger (2006), Algorithm theoretical basis document for GRACE level-1B data processing V1. 2. Jet Propulsion Laboratory, California Institute of Technology.

Xie, J., (2005), Master Thesis, Implementation of parallel least squares algorithm for gravity field inversion, Dept. of Civil & Environmental Engineering & Geodetic Science, Ohio State University.

Xie, J., S. Ge, S. Han and C. Shum (2004), Implementation of parallel least square algorithm for gravity field inversion, *J. of Geospatial Eng.*, 6(1), 47–57.

Yamamoto, K., T. Otsubo, T. Kubo-Oka, and Y. Fukuda (2005), A simulation study of effects of GRACE orbit decay on the gravity field recovery. *Earth, planets and space*, 57(4), 291-295.

Yi, S., and W. Sun (2014), Evaluation of glacier changes in high-mountain Asia based on 10 year GRACE RL05 models, *J. Geophys. Res. Solid Earth*, 119, 2504–2517, doi:10.1002/2013JB010860.

Yi, W (2012), The Earth's gravity field from GOCE, *PhD diss.*, München, Technische Universität München.

Zeng, Y.Y., J.Y. Guo, K. Shang, C. Shum, and J.H. Yu, On the formulation of gravitational potential difference between the GRACE satellites based on energy integral in Earth fixed frame, *Geophys. J. Int.*, 2015.

Zhao, Q., J. Guo, Z. Hu, C. Shi, J. Liu, H. Cai, and X. Liu (2011), GRACE gravity field modeling with an investigation on correlation between nuisance parameters and gravity field coefficients. *Adv. Space Res.*, 47(10), 1833-1850.

Zhou, B., et al. (2011), The great 2008 Chinese ice storm: its socioeconomic-ecological impact and sustainability lessons learned, *Bulletin of the American Meteorological Society*, 92.1 (2011): 47-60.